
**Reviewer Report for the Center for Independent Experts on the SEDAR Gulf
of Mexico Data-Limited Species Review Workshop (SEDAR 49) held
November 1-3, 2016, in Miami, Florida**

Prepared for:

Center for Independent Experts

By:

Jamie Gibson
100 Allison Coldwell Road,
Wolfville, Nova Scotia, Canada
B4P 2R1

Executive Summary

This document forms my independent reviewer report of review activities and findings for the 49th Southeast Data, Assessment and Review (SEDAR 49) Review Workshop, held November 1-3, 2016 in Miami, Florida. Eight data-limited species in the Gulf of Mexico, including: Red Drum, Lane Snapper, Wenchman, Yellowmouth Grouper, Speckled Hind, Snowy Grouper, Almaco Jack and Lesser Amberjack, were assessed during SEDAR 49.

Overall, the Data Workshop Report and supporting working papers clearly presented the information available for assessing these stocks. Data available for the assessment fall into the broad categories of delineation of stocks, life history parameters, commercial fishery statistics, recreational fishery, total removals, measures of fishing effort, fishery-dependent and fishery-independent indices of population abundance, length-frequency data, age-frequency data, as well as other information (reference periods, depletion levels). This material was very well-organized in the workshop reports, background information and presentations, which facilitated the review.

The primary analytical method used in SEDAR 49 is the Data-Limited Methods Toolkit. The toolkit uses simulations to evaluate the performance of multiple data-limited approaches via management strategy evaluation (MSE). The data-limited methods provide an evaluation of the potential management procedures given the data available for each species. Once potential methods are chosen, catch recommendations are developed that are appropriate for each method. I consider these methods to be sound, to be consistent with best practices in conservation biology, and highly appropriate given the data available for the species assessed in SEDAR 49.

As applied in SEDAR 49, the output from these methods are options for management procedures rather than the estimates of stock size, fishing mortality rates, and reference points produced by more traditional assessments. The work undertaken during SEDAR 49 provides a thorough review of potential management procedures and operating models that could be utilized for these species.

If these methods are to form the basis for future assessments for these species, I believe that, at least initially, research towards tailoring the data-limited methods used in this assessment specifically for the provision of catch advice, and specifically to the requirements for the individual species, would be expected to lead to greater improvements in these assessments than improving data for these species.

Table of Contents

Executive Summary	1
1.0. Background	3
2.0. Individual Reviewer Activities	3
3.0. Summary of Findings in Accordance with the TOR's.....	4
4.0. Conclusions and Recommendations in Accordance with the TOR's	22
5.0. References.....	28
6.0. Appendices.....	30

1.0. Background

This document forms my independent reviewer report of review activities and findings for the 49th Southeast Data, Assessment and Review (SEDAR 49) Review Workshop, held November 1-3, 2016 in Miami, Florida. Eight data-limited species in Gulf of Mexico, including: Red Drum, Lane Snapper, Wenchman, Yellowmouth Grouper, Speckled Hind, Snowy Grouper, Almaco Jack and Lesser Amberjack, were assessed during SEDAR 49. The Data-Limited Methods Toolkit (DLMtool; Carruthers et al. 2014, Carruthers et al. 2015, Carruthers and Hordyk 2016), was the primary tool used to assess these stocks. SEDAR 49 consisted of a data workshop, an assessment workshop carried out via webinars, and the peer-review workshop. Working papers, background material, and the data workshop and assessment reports (Appendix 1) were provided to the Review Panel (Appendix 3) two weeks prior to the Review Workshop. Together, these documents provided a comprehensive and clearly presented compilation of the data available for these species, the decisions that were made about how to use these data, the analyses that were undertaken, and the results and conclusions of the assessment. These documents provided an excellent basis upon which to base the peer review. The analytical team did a lot of work after the Review Workshop and provided the results of multiple sensitivity analyses as an addendum after the Review Workshop. These results were very helpful for understanding the model behavior.

2.0. Individual Reviewer Activities

I conducted my activities in accordance with the Statement of Work (SOW) provided for this review (Appendix 2). The terms of reference (TOR's) for both the Review Workshop and this individual report are contained in the SOW. Prior to the meeting, I reviewed all the assessment and background documents provided for the workshop. I participated in the Review Workshop in Miami, Florida, November 1-3, 2016. During the meeting, I actively participated as member of the Review Panel, and questioned and discussed several aspects of the data and models. The meeting was relatively informal, with a lot of discussion during the presentations. The Analytical Team openly discussed the work undertaken for this assessment and they were timely in their responses to questions and requests for further analyses. Because of their efforts, I believe the review workshop was a very productive meeting. After the Review Workshop, I prepared this individual, independent report and assisted with writing the Review Workshop Report. As outlined in Appendix 2, this independent report is intended to summarize review activities completed during the panel review meeting, including providing a detailed summary of findings, conclusions, and recommendations for each TOR. The following sections in this document contain my personal perspectives about this assessment.

3.0. Summary of Findings in Accordance with the TOR's

TOR 1. Review any changes in data following the Data/Assessment workshop and any analyses suggested by the workshop. Summarize data as used in each assessment model. Provide justification for any deviations from Data/Assessment Workshop recommendations.

The data used during SEDAR 49 followed the recommendations of the Data and Assessment Workshops. The Review Panel did request some sensitivity analyses using different values, but these were intended as exploratory, not recommendations to change the data. Data used in the assessment fall into the broad categories of delineation of stocks, life history (including age and growth data, natural mortality, reproduction, meristic conversions and steepness), commercial fishery statistics (including commercial landings, discards and bycatch), recreational fishery statistics (including landings and discards), total removals, measures of fishing effort, indices of population abundance (including fishery-dependent and fishery-independent surveys), length-frequency data, age-frequency data and other information relating to topics such as reference periods, depletion levels, species misidentification, fishery characteristics, discard mortality and size of discards. In general, I think the presentation of this material was very well-organized in the workshop reports, background information and presentations, which facilitated the review.

Delineation of Stocks

The Analytical Team provided a summary of the genetic analyses and other information available for identifying stocks for the species assessed in SEDAR 49. With the exceptions of Red Drum and Lane Snapper, for which there is evidence of genetic divergence within the GOM, there is relatively little information available for delineating stocks for any of the eight species. A single stock was assumed for each species.

Life History Information

The Analytical Team very thoroughly evaluated the strengths and weaknesses of the life history information available for use in SEDAR 49. A review of the peer-reviewed literature (Adams et al. 2016), and published and unpublished reports, was used to identify sources of life history information; and a scoring reliability rubric, based on sampling considerations (number of samples, temporal duration and frequency, spatial aspects, method, etc.), the quality of the data collection and analysis, and the overall reliability of the work, was created as a semi-quantitative method to score the overall utility of the work for providing means and ranges of values for life history parameters for use in the DLM tool. This approach provided a thorough and transparent process for providing life history information for the models. Additional analyses of life history data were undertaken with stock specific data where possible.

In addition to the summaries provided in the Data Workshop Report, the Review Panel was provided with reports summarizing the operating model inputs for each species. These provided a very useful summary of the sources of life history information for each species.

Information about age and growth was compiled via a literature review and/or analyses of data undertaken as part of SEDAR 49. The three parameter von Bertalanffy (VB) growth equation

was used for all species. Red Drum was the most data rich species, including information from six growth studies, several assessments, and five datasets containing over 8,000 age estimates. Growth information was not available for Almaco Jack or Lesser Amberjack, and VB growth parameters derived from Greater Amberjack parameters were adjusted for size differences between the species. Quality of the data was variable for the other species. For example, as pointed out by the Analytical Team, the single study of growth for Wenchman was based on 115 specimens with a maximum size much smaller than the largest specimens collected from the GOM.

Information about the age and length at maturity for these species was primarily compiled from the literature review. Maturity information was not available for Wenchman, Lesser Amberjack or Almaco Jack. For these species, maturity information was extrapolated from information available for closely related species.

Length-weight conversion factors were developed for all species using species-specific data obtained from multiple fishery-dependent and fishery-independent sources.

For all species, instantaneous natural mortality rates were estimated from longevity using the updated Hoenig equation (Then et al. 2015). This decision was made based on a review of empirical methods for estimating natural mortality (SEDAR 2015). Estimates of maximum age were taken from the literature review and required judgement by the Analytical Team to select an appropriate age. For example, the value used for Speckled Hind was older than the value reported in the two available studies based on the results of radiocarbon data.

Very little information was available about the steepness of the stock-recruitment relationship for any of the species assessed during SEDAR 49. Based on the information in the literature review, the Assessment Workshop recommended values used in other stock-specific assessments (Red Drum), assessments of congeners (Yellowmouth Grouper, Speckled Hind, Lesser Amberjack and Almaco Jack), or from meta-analyses (Lane Snapper, Wenchman and Snowy Grouper).

Commercial Fishery Statistics

Data used to construct total landings for the commercial fisheries are housed in NOAA's Southeast Fisheries Science Center's Accumulated Landings System (ALS). The terminal year of 2014 was the same for all species, but the initial year of data used in the model varied from 1962 for Red Drum to 1991 for Lesser Amberjack and Almaco Jack. The selection of the initial year was based on judgment of the data quality in terms of its completeness, the potential for species misidentification or issues with the assignment of unclassified fish. Uncertainty in the landings was estimated using expert opinion.

The Data Workshop and Assessment Workshop reports provided estimates of discard rates for the commercial vertical line, bottom longline and shrimp trawl fisheries. Discards were considered negligible for some species/fishery combinations (e.g., Red Drum, Yellowmouth Grouper, Lesser Amberjack and Almaco Jack in the three fisheries), but it was considered more significant in other combinations (e.g., Lane Snapper and Wenchman in the shrimp fishery; Snowy Grouper and Speckled Hind in the vertical line and bottom longline fisheries). For the most part, discard mortality rates were assumed based on experience with other species.

Recreational Landings

Recreational landings were obtained from four separate sampling programs: Marine Recreational Fisheries Statistics Survey (MRFSS) and the Marine Recreational Information Program (MRIP); the Southeast Region Headboat Survey (SRHS); the Texas Parks and Wildlife Department (TPWD); and the Louisiana Creel Survey. Several adjustments and modifications to the surveys have been made to produce more reliable estimates of the landings and bycatch. Annual removals associated with discarding were calculated. Dead discard biomasses were estimated by multiplying the number of discards by the discard mortality rate and the average weight of discarded fish.

Total Removals

Total removals were calculated as the sum of the commercial landings, commercial dead discards, recreational landings and recreational dead discards. The DLM toolbox requires a single estimate of the uncertainty associated with the total removals. This was calculated using the CVs associated with each of the four categories.

Representative Fleets and Fishing Effort

A single representative fleet and an estimate of fishing effort are required for each species in order to use the DLM toolkit. The Analytical Team selected the fishery with the highest landings as the representative fleet. Commercial fisheries were considered most representative for Speckled Hind (bottom longline), Snowy Grouper (bottom longline), Lesser Amberjack (vertical line), and Wenchman (finfish trawl). Recreational fisheries were considered most representative for Red Drum, Lane Snapper, Almaco Jack and Yellowmouth Grouper. The number of fishing trips within a year was used as the metric for fishing effort in both the recreational and commercial fisheries. The toolkit also requires an estimate of the selectivity of the fishery as an input. The age-at-first-capture and the age-at-full-selectivity and their associated uncertainty were derived from length-frequency data for the representative fishery and information provided by the fishermen.

Indices of Abundance

The Data Workshop Report and background documents provided thorough descriptions of the available abundance indices, methods used for their calculation, their strengths and weaknesses, and justification for their inclusion or exclusion as data inputs for the model. Abundance indices were available from both fishery-independent and fishery-dependent sources. Fishery-independent surveys that were considered for inclusion in the model are: the SEAMAP Summer Groundfish Survey, the Southeast Fisheries Science Center Mississippi Laboratories (MSLABS) Small Pelagics Survey, the SEAMAP Reef Fish Video Survey, the NMFS Panama City Laboratory Trap and Camera Survey, and the Dauphin Island Sea Lab (DISL) Bottom Longline Survey. Fishery-dependent data sources that were considered are: the Headboat Survey, the Marine Recreational Fisheries Statistics Survey / Marine Recreational Information Program, and the NMFS Gulf of Mexico Reef Fish Logbook Program.

Several criteria were used when selecting a single survey for inclusion in the model, including: the temporal and spatial coverage, the number of samples, consistency with other indices, the extent to which it samples the population and the proportion of positive catches. The Headboat Survey was selected to provide an abundance index for Lane Snapper, the DISL Bottom

Longline Survey was selected for Red Drum, the MSLABS Small Pelagics Survey for Wenchman, and the SEAMAP Reef Fish Video Survey was selected to provide abundance indices for Lesser Amberjack, Almaco Jack, Snowy Grouper and Yellowmouth Grouper.

No abundance indices were recommended for use in SEDAR 49 for Snowy Grouper or Speckled Hind. Although both species are detected in the SEAMAP reef fish video survey, the Analytical Team concluded that neither is detected frequently enough to provide a reliable index.

Depletion Levels

An estimate of the current level of depletion of the stock is required as an input for the DLMtool. The only species for which an estimate of depletion was available was Red Drum, which came from an assessment of Red Drum in Florida waters. Depletion estimates for the other species were derived from other, similar stocks that have been assessed. Current stock depletion can be estimated within the DLMtool. The estimates are highly uncertain (as is discussed in the Assessment Workshop Report), but they do provide another source of information about depletion levels. The Analytical Team used the estimates from the DLMtool to update the values used for depletion in the base model. Additionally, the Analytical Team thoroughly investigated the assumptions about depletion levels via sensitivity analyses.

Reference Period

To program and evaluate the current approach implemented by the Gulf of Mexico Fishery Council, reference time periods provided in GMFMC (2011) were used for seven of the eight species. A reference period for the eighth species, Red Drum, was chosen at the Assessment Workshop. The reference period also establishes baseline values for indices that are used in some of the harvest control rules.

Length Frequency Data

Length frequency data was obtained from both fishery-dependent and fishery-independent sources. A port sampling program (the NOAA Fisheries Southeast Fisheries Science Centre Trip Interviewer Program (TIP)) and logbook reporting were the sources of length frequency data for the commercial landings. Length frequency data for the recreational fisheries were obtained from several sources, including the Marine Recreational Fisheries Statistics Survey (i.e., the Marine Recreational Information Program, MRIP), the Head Boat Survey, the Texas Parks and Wildlife Department database (TPWD), the Florida Fish and Wildlife Conservation Commission (FWC), the Gulf States Marine Fisheries Commission FIN database (GFIN), and the TIP database. Fishery-independent sources of length frequency data included the NMFS small pelagics survey, the SEAMAP groundfish survey, the SEAMAP reef video survey, the Panama City video survey, and the Panama City trap survey. Sample sizes for some species encountered in these surveys were too small for analysis.

With respect to the data used in the assessment, no specific changes were recommended by the Review Panel, although sensitivity analyses were requested as diagnostics or to help the panel better understand the workings of the model. With respect to data inputs, sensitivity analyses pertaining to the “ L_{inf} ” growth parameter and the stock-recruitment steepness parameter (h) were undertaken by the analytical team and presented or provided to the review panel.

TOR 2. 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?

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?

In general, I believe that the data decisions made by the Data Workshop and Assessment Workshop are sound and robust. The uncertainties in the data were very well documented in the Data Workshop and Assessment Workshop reports and are as expected given these are data-limited species. I believe the data are properly applied within the assessment model and that the input data series are sufficient to support the general approach used in this assessment. However, there are considerable uncertainties in the data particularly with respect to life history parameters, and in some cases results are sensitive to the values assumed in different model runs. My opinions about the strengths and weaknesses of the data sources and decisions follow below.

Delineation of Stocks

For stock assessment purposes, a practical definition of a stock, consistent with the goals of an assessment (to determine whether mortality rates and abundances are within acceptable limits) is “a population or subpopulation of a particular species of fish, for which intrinsic factors (growth, reproductive rates, carrying capacity, natural mortality and mortality caused by human activities) are the significant factors determining the stock's dynamics, and for which extrinsic factors (immigration and emigration) are minimal and can be ignored.” If a single stock is assumed for a fishery harvesting from more than one stock, the risks include that stocks with lower productivity would be impacted to a higher degree, and that if fishing effort is concentrated on a small number of isolated stocks, the removal rates for these isolated stocks may not be sustainable.

Where genetic analyses do show significant divergence among groups of fish, it can be interpreted as evidence of an underlying stock structure; however, the absence of significant divergence does not provide evidence of only a single stock because populations may not have been isolated long enough for divergence to have occurred or stray rates may be high enough to homogenize the gene pools even if the populations are essentially demographically uncoupled. As summarized in the Data Workshop Report, there is evidence of genetic divergence in the northern GOM for Red Drum, and for Lane Snapper there is evidence from microsatellite data for two stocks, but specific populations have not been delineated for either species.

Overall, I think the decision by the Analytical Team to model each species assessed in SEDAR 49 as a single stock unit is a practical one, given both the data limitations and the difficulties of splitting data on the scale of smaller potential stocks (e.g., assigning landings to individual stocks in a mixed stock fishery). This is a common assumption in fishery assessments. However, if fishing effort is distributed over a small geographic area, using a smaller potential stock might help reduce risks associated with the stock delineation decisions.

Life History Information

Life history parameter values used in this assessment are mainly derived from the literature and are incorporated into the assessment by setting minimum and maximum values and randomly drawing values from a uniform distribution defined by these bounds. While the range of values for an individual parameter might be appropriate, the overall productivity of a population is determined by the set of values drawn for all parameters. Some combinations of parameters might be biologically unrealistic, but life history parameter covariance is quite difficult to incorporate into the simulations. One step towards reducing parameter covariance would be to choose parameters that are not dependent on other model parameters. For example, the steepness parameter in the stock-recruitment (SR) relationship depends on the slope at the origin of the SR relationship, but also on the natural mortality rates, growth parameters, maturity parameters, and length-weight conversion parameters, most of which were assumed uncorrelated in SEDAR 49. This specific issue would be addressed if the SR relationship was parameterized in terms of the (biologically meaningful) slope at the origin and an abundance scalar such as the asymptotic recruitment level (Myers et al. 2001). Additionally, metrics that are functions of several life history parameters, such as steepness, $SPR_{F=0}$, the maximum lifetime reproductive rate or other similar metrics could be calculated and could provide a mechanism for filtering out combinations of parameter values that are biologically unrealistic, if limit values for these other metrics were included in the model.

As shown in the Assessment Workshop summary of the data reliability (Assessment Workshop Report Table 2.1), the stock recruitment steepness parameter is one of the life history parameters for which there is the least information. I endorse the approach of using values from the literature and from other assessments, but I think it is important to distinguish between values estimated from data and values assumed in other assessments. The distribution for steepness used for Red Drum had an upper value of 1.0, based on values assumed in SEDAR 44. This value was likely used to estimate a mean recruitment and deviates rather than being based on a belief that the value was plausible. Additionally, there are differences in the stock recruitment model used in the Myers et al. (1999) meta-analysis (a Ricker model) and the model used in SEDAR 49 (a Beverton-Holt model). While the effect of this choice is unclear, it may be slightly precautionary. There are other models that could potentially be used in a meta-analysis, such as the hockey-stick (Barrowman et al. 2001), that could potentially help with determining the range of potential values for the slope at the origin of stock recruitment relationship. Finally, the steepness values from the meta-analysis may not be directly applicable if values for natural mortality, growth and maturity differ between the meta-analysis and those assumed in SEDAR 49.

To help the Review Panel better understand the forward projecting model in the DLM tool, the Analytical Team ran the model with the steepness fixed at the lower bound and provided these results in the addendum. To me, it is not clear how to interpret some of these results. For example, for Red Drum, the long term yield when values for steepness are drawn from a uniform distribution with bounds of 0.8 and 1.0 are lower than the long term yield when the steepness is fixed at 0.8 (Islope0 results of 12.7 versus 21.2). These results seem counter-intuitive but are not evident for all species. Other comparisons are more as expected. A general reduction in the probability of not overfishing was evident when a lower steepness was assumed, but this was not evident in all cases. For example, for the Lane Snapper using the LstepCC0 model, the

probability of not overfishing went up when a lower steepness was assumed, but with lower short-term and long-term yields. The results of the comparison do show the sensitivity of both the operating model selection and other results to the assumed value for steepness.

Commercial Fishery Statistics, Recreational Landings, and Total Removals

From my perspective, SEDAR 49 provided a thorough overview of the removals associated with the commercial and recreational fisheries. The reports provided good justifications for decisions about how to include these data, such as the length of the time series to use in the model.

Although the CVs for the total removals contain an element of subjectivity (expert opinion was used to estimate the uncertainty associated with the commercial landings), I think this is a minor issue given some of the other uncertainties and, if needed, could be evaluated using sensitivity analyses.

Representative Fleets and Fishing Effort

Overall, I found the decisions about the representative fleets and fishing effort to be sound and well documented. Selecting the fishery with the highest proportion of the total removals as the representative fleet is practical, as is the decision to use the number of fishing trips within a year as the metric for fishing effort, given the data. Decisions about the age-at-first-capture and the age-at-full-selectivity, as well as their associated uncertainty were derived from the available data, but also included information provided by the fishermen to help ensure that the decisions were sound.

Indices of Abundance

Fishery-independent and fishery-dependent abundance indices were thoroughly reviewed in the Data Workshop Report and background documents. In situations where more than one index was available, decisions about which index to include in the model were well described. However, this is a subjective decision. The Review Workshop did include consistency with other indices as one of the criteria for selecting an index; however, if indices diverge, sensitivity analyses might help evaluate the effects of this decision on the resulting catch advice (although this approach might make it difficult to choose a management procedure for future use).

As discussed in the Data Workshop Report, there is the potential to develop indices for both Snowy Grouper and Speckled Hind using data from the commercial logbook program, although further work is needed to determine whether abundance changes in the commercial data can be separated from fleet dynamics or fishing behavior. This is more of a topic for future research, although if it is possible to do so, development of a fishery independent survey might better reflect future changes in the status of these species.

Depletion Levels

As was well described in the workshop reports, the levels of depletion are a source of uncertainty in this assessment. I agree with the Review Panel comments that using depletion levels for well-studied species that are targeted may not be appropriate for bycatch species. However, the Analytical Team did undertake sensitivity analyses for the assumed level of depletion and this uncertainty was carried forward throughout the assessment via these analyses.

Reference Period

The reference period had two main functions in SEDAR 49: it was used to evaluate current practices, and also to provide a baseline status associated with the index values during that time period. As such, it can become an important input. During the meeting, the Analytical Team explained that one of the criteria used to select the reference period was to choose a period with reasonably constant catch and I agree with the Review Panel Report that there are arguments for and against this approach. However, particularly for harvest control rules that use indices, it might be more important to use a period when the depletion level may be best known. For the species assessed in SEDAR 49, depletion is not well known. The Analytical Team did sensitivity analyses for the level of depletion with variable results among species. These sensitivity analyses may partly address the uncertainty associated with the reference period.

Length Frequency Data

I do not have any specific comments about the length-frequency data or concerns about its use in this assessment. The data and decisions about its use are well described and appear sound.

TOR 3. 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 multi-model approach was used to assess the eight stocks in SEDAR 49. The primary analytical method was the Data-Limited Methods Toolkit (DLMtool; Carruthers et al. 2014, Carruthers et al 2015, and Carruthers and Hordyk 2016). This program uses simulations to evaluate the performance of multiple data-limited approaches via management strategy evaluation (MSE). Once potential methods are chosen, catch recommendations are developed that are appropriate for each method. The second analytical approach was the application of a mean length-based mortality estimator assuming non-equilibrium conditions to estimate the total mortality rate (Gedamke and Hoenig 2006). The third analytical approach, catch curve analysis, was also used to estimate total mortality. This later method was used only for red drum. Appropriately, the greatest emphasis was placed on the data-limited tools.

The catch curve analyses were carried out via least-squares regression on the log-transformed numbers-at-age, where the slope of the regression provides the value for the instantaneous total mortality rate (Z). This is a commonly used approach, but may not provide the best estimate of Z . Catch curve analysis has a long history of use for estimating mortality in fisheries biology and the methods for using age-frequency data continue to be studied and improved (e.g. Smith et al. 2012, Millar 2015). Assumptions underlying the analyses include (Smith et al. 2012): recruitment is constant or at least varies without trend; mortality is constant over time and age classes; above some age (which can be difficult to identify), all animals are equally vulnerable to the fishery and the sampling process; and there are no errors in the estimation of age composition. Both Smith et al. (2012) and Millar (2015) tested several methods of fitting catch curves using simulated data. Smith et al. (2012) found that other methods outperformed least squares regression, and Millar (2015) developed a mixed effects model with random effects on age that helps address the issue of recruitment variability; however, no one method has been demonstrated to be optimal in all situations. However, because the primary use of the total mortality estimates for red drum was to inform estimates of natural mortality used in the DLM

tool, and because the sensitivity analyses were done for this variable, I do not believe the use of a different method would lead to a change in the assessment results.

The length-based mortality estimator is an assumption rich method for estimating the total mortality rate from the von Bertalanffy growth model parameters, an estimate of the length at full selection and the mean length of fish greater than the length at full selection. The approach has been extended to be applicable in non-equilibrium conditions (Gedamke and Hoenig 2006). The Assessment Workshop did attempt to fit this model for each of the species assessed during SEDAR 49, but did not obtain satisfactory results for any of the species, because of concerns about the violation of the assumptions, lack of growth information, or implausible results. The Assessment Workshop did not place emphasis on these results and appropriately focused on the results of the analyses using the DLM tool in the Workshop report. The remainder of this review is focused on the application of this method.

a) Are the data-limited methods scientifically sound and robust?

The Data-Limited Methods Toolkit (DLMtool; Carruthers et al. 2014, Carruthers et al. 2015, and Carruthers and Hordyk 2016) was the primary analytical tool used in SEDAR 49. This tool has been used in other studies and has been peer reviewed in the primary literature. MSE approaches have been developed in the scientific literature. I agree with the statement in the Review Panel Report that the equations are, for the most part, standard equations that have been used extensively in the past. I consider these methods to be sound.

In many ways, the population simulations undertaken for SEDAR 49 are more similar to the population viability analyses (PVAs) used in conservation biology than they are traditional stock assessment models. PVAs are used extensively in conservation biology to predict both the risk of extinction for populations and species, and to evaluate management strategies to recover at-risk populations. In a PVA, a population dynamics model is used to determine how the probability of persistence is affected by current conditions and future perturbations (Beissinger and McCullough 2002), and models are often used to evaluate how future management actions or environmental changes may influence the probabilities of extinction or of achieving recovery goals (Reed et al. 2002). PVAs are well studied.

A criticism of PVAs is that their longer-term abundance predictions are highly uncertain (e.g., Taylor 1995; McCarthy et al. 1996; Ludwig 1999). For this reason, some authors have suggested that the best use of PVAs is to assess relative risk among a set of possible management actions (Akçakaya and Raphael 1998; Beissinger and Westphal 1998; Lindenmayer and Possingham 1996, McCarthy et al. 2001). For example, McCarthy et al. (2003) used a simulation study to test preferred management strategies and found that they were able to identify the better of two management strategies 67–74% of the time using 10 years of data, and 92–93% of the time with 100 years of data. While their results would not specifically apply to this assessment because of differences in the data, I do believe the use of simulation models to evaluate alternative management procedures, as was done in SEDAR 49, is consistent with best practices in conservation biology.

Notwithstanding this endorsement of the approaches used in SEDAR 49, more complex models would be expected to produce different results. For example, MacCall (2002) used a two species model with low frequency environmental variability to show how rebuilding prospects and fishery management differed under different environmental regimes. While I'm not suggesting more complex models would be more appropriate, the example does show that with respect to ensuring the goals of fishery management are met, computer simulations cannot replace data. However, the MSE approaches do provide excellent guidance on how to make the best use of the limited data available for some species.

b) Are the methods appropriate given the available data?

Long-term population simulations, such as PVAs or those carried out with the DLM tools, allow one to explore the logical implications of current knowledge and assumptions and are therefore highly appropriate given the data available for the species assessed in SEDAR 49.

c) Are the data-limited models configured properly and used in a manner consistent with standard practices?

The Assessment Workshop provided working papers summarizing the operating model parameters for the DLMtool as configured for each species. These working papers made it easy to check the assessment inputs. Overall, I agree with the Review Panel Report that the setup appears to match the recommendations from the Data Workshop.

The Analytical Team did develop some code specifically for use in SEDAR 49. During the Review Workshop, I questioned whether the equations in Section 3.1 of the SEDAR 49 Addendum match those in Geromont and Butterworth (2014) corresponding to Figure 2 from their supplementary material (reproduced on page 7 of the Addendum), specifically with respect to the catch recommendation smoothing parameter w . The analytical team responded that the default value, $w=0.5$, was hardwired in the code as written in the equations in the Addendum. This approach would work for the default values, but would not for values of w other than 0.5 and would not reproduce the figure on page 7 of the Addendum. This code could be checked to ensure it matches the equations in Geromont and Butterworth (2014). Further checking of the code would help to ensure that other inconsistencies are not influencing the results of the analyses.

As discussed under TOR 2, parameter covariance could result in some simulated abundance trajectories that are not realistic. Further development of the models to reduce the effects of potentially unrealistic parameter combinations would be expected to improve the results from the models used in SEDAR 49.

Rather than using a single value of each life history parameter, values are input as uniformly distributed variables with an upper and lower bound. As implemented in SEDAR 49, single methods were used to estimate natural mortality, and when values were taken from meta-analyses, the meta-analyses typically employed a single model to estimate the parameter and its variance. The use of more than one model to estimate the parameters could result in additional support for the range of plausible values (assuming different values were obtained).

Further development of visual aids to guide interpretation of analytical results would be beneficial. One approach would be to add more information to the figures summarizing the simulation results. Results were shown using the mean and 5th and 95th percentiles of the simulation runs. Adding additional percentiles to the same plot (possibly the 25th and 75th) would better allow for visualization of the skewness of the distribution around the mean values. Display of single population projections also allows for a qualitative assessment for the realism of the simulations (e.g., are the simulated abundance changes similar to those seen for related species, are abundance increases realistic given what is known about the biology of the species).

In the results displayed during SEDAR 49, some trajectories appeared to drop to quite low abundances and recover. This may or may not be realistic depending on whether there is a feedback mechanism that would reduce human-induced mortality as abundance declines, which might not be the case for bycatch species. Inclusion of a quasi-extinction threshold, as is typically done in PVAs, might improve the realism of the simulations.

Based on experience with PVAs, temporal autocorrelation in model parameters (e.g., natural mortality rates, recruitment autocorrelation) can be a key determinant of the probability of meeting management objectives. However, given the difficulties of estimating parameter variances, estimating parameter autocorrelation will remain problematic. This topic was not explored during the Review Workshop, but could warrant exploration as these models are further developed via sensitivity analyses with respect to assumptions about temporal autocorrelation.

d) Are the quantitative estimates produced reliable? Does the method produce management metrics (e.g. OFL, ABC) or other indicators (e.g. trends in F or Z, probability of overfishing) that may be used to inform managers about stock trends and conditions?

As applied in SEDAR 49, the data-limited methods provide an evaluation of the potential management procedures given the data available for each species. The output from these methods are options for management procedures rather than the estimates of stock size, fishing mortality rates, and reference points produced by more traditional assessments. Although probability densities for catch recommendations are produced as part of the evaluation, as discussed under TOR 4, it is not clear that they capture the full range of catch recommendations. Applying these results in terms of catch recommendations could require adopting a procedure, and as recommended in the Workshop Report, fine tuning the model to ensure that index scalars, assessment frequency and other control variables in the operating model are set appropriately for the species. The compilation of life history information for the species assessed during SEDAR 49, the synopsis of the data available that can be used to evaluate management procedures for the species and the application of the data-limited tools to evaluate potential management procedures is an excellent step in this direction.

The uncertainty in how to use the evaluations of potential management procedures within the current paradigm was evident in the presentation of the catch recommendations. For several species, more than one management procedure met the performance criteria, while at the same time producing probability distributions for the catch recommendations associated with the

procedure. The Assessment Workshop addressed this issue by producing a joint distribution that was recommended as a basis for providing management advice. This approach appeared to result from uncertainty about future management procedures and might not produce the best catch recommendation if a single procedure or different procedure is adopted in the future. Developing a catch recommendation via specific evaluation of the management procedure to be used would be expected to result in an improved recommendation.

A clear interpretation of the how the catch recommendations from the DLM analyses relate to OFL, ACL, ABC, ACT or some other metric was not available at SEDAR 49. Particularly given that a probability distribution is produced for the recommendation, the potential to use different percentiles from the distribution for different metrics could be explored. The interpretation of the output might also be situation-specific, and might differ among populations.

TOR 4. Consider how uncertainties in the assessment, and their potential consequences, are addressed.

Because the species assessed in SEDAR 49 are data limited, uncertainty in the data sources and assessment results is relatively high. Sources of uncertainty arise in the data inputs, in the selection of life history parameter values, in setting up the operating model and characterizing the fishery (e.g., selectivity), in assigning the state of depletion, as well as in observing the population in the future. Additionally, because the use of the DLM tool for the provision of catch advice is relatively new, further uncertainty arises from how best to use the tool for catch advice.

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.

As described under TOR 1, the Analytical Team fully considered potential data inputs, conducted a literature review and created a scoring reliability rubric to choose ranges of life history parameter values. Uncertainty in the life history parameters, fishery selectivity, and further observations of the population (future index values) were incorporated into the simulations using Monte Carlo methods. These methods are highly appropriate.

In addition, in using the DLM tool the team undertook a large number of sensitivity analyses to evaluate the effects of the decisions about depletion levels, values for steepness, hyperdepletion or hyperstability in the indices, annual variability in growth parameters, assessment frequency, CVs for future indices, lower values for steepness, and for Red Drum, different values for steepness, and different thresholds for index values used in the catch rules for the methods that incorporated indices. I found these analyses very useful for evaluating the model runs and understanding the uncertainty in the model runs.

The model output includes a probability distribution for the catch recommendation associated with each potential operating model. Extremely low (near zero) catches would be expected to meet the performance criteria established for SEDAR 49 (optimizing landings with respect to MSY was not a criterion), but were assigned low probability in the probability density for the catch recommendations. Intuitively, a constant catch scenario should always be able to meet the

performance criteria for SEDAR 49 if the catch is set low enough. The Analytical Team explained that this was due to constraints within the model that limited how much the catch recommendation could change from recent landings. For this reason, I am not sure that the model, as configured, could fully explore the possible range of catch recommendations. If not, both the uncertainty associated with the catch recommendation and potentially its mode may be conditional on this constraint, which may or may not be the “optimal” recommendation for each species if recent catch levels are low or high. A broader range of catch recommendations for each operating model might increase the number of options available for operating models. The effect of constraints on the catch recommendations from a single operating model was not fully explored during SEDAR 49, but is necessary to be able to interpret the probability distributions for the catch recommendations.

As used for SEDAR 49, the DLM tool was set up primarily for evaluating potential management procedures and operating models, and not for evaluating whether a specific catch recommendation would meet the performance metrics (there is uncertainty associated with the catch recommendation). An additional step, involving feeding the specific catch recommendation back into the operating model to ensure that performance metrics are met given the uncertainty in the operating model input parameters (use different random values) would address this source of uncertainty. Sensitivity analyses with respect to key assumptions (e.g., depletion levels) at this stage would be expected to further reduce uncertainty in the catch recommendation.

As described above, the long-term population projections are in many ways analogous to population viability analyses used in conservation biology. The method of calculating the probability of meeting performance criteria used in SEDAR 49 differs from the approach typically used in PVAs. As implemented in SEDAR 49, the probability of meeting the performance metrics was calculated across all years and simulations simultaneously. However, each simulation is a potential realization of future conditions that either meets management objectives (stated as performance criteria) or does not. Within a PVA, each simulated population is scored based on whether it meets the evaluation criterion (e.g., the population recovers or it goes extinct). The probability of meeting the objective is calculated from the proportion of simulated population trajectories that meet the criterion. In this system, performance metrics and standards, based on management goals, are two-tiered, including criteria that are applied to each individual simulation to determine whether it meets the metric or standard, as well as risk acceptance criteria applied across simulations based on the probability that the standard is met. Examples of performance metrics that could be applied within a single simulated trajectory include: the proportion of the years during which the population is overfished; the proportion of the years the population is in an overfished state; the proportion of the years that the population is above or below some abundance threshold; or, in the case of rebuilding, whether a simulated population meets rebuilding objectives within a specified time. Each simulated population trajectory either meets the performance metric, or does not. The probability of meeting the objective can then be calculated as the proportion of simulated populations that meet the performance metric. This probability can then be compared with the risk tolerance criterion for the objective. The beauty in using these kinds of simulations is that the performance metrics can be very explicitly tied to the management objectives for the species and the relative probability of meeting the objectives using various procedures can be used to guide future decisions.

With respect to uncertainty arising from decisions about fine tuning and presenting results from a new tool, the management strategy evaluation simulations do provide the opportunity to set up the model to align with the management needs for each species. These needs might be expected to vary depending on factors such as whether the species is taken primarily as bycatch or is being targeted and is being managed for MSY; or whether the population is in a depleted state and whether rebuilding is being considered, or other factors. Further, during SEDAR 49, there was uncertainty about how the catch recommendations were to be interpreted. The results from SEDAR 49 certainly do show the range of operating models and management procedures that are available for these species, but it was less clear how to choose the model and procedure. Additionally, while the performance metrics used to choose candidate management procedures may be appropriate for bycatch species, they were not configured in a way to attempt to optimize yield, as would be preferable for a targeted species. Refinement of the questions to be addressed via the simulation model, by first stating the goals, then establishing a set of performance metrics related to these goals, and then establishing risk tolerance criteria for each of the performance metrics would aid in reducing the uncertainty associated with the provision of advice associated with the use of this new tool. The comments in this paragraph are not intended in any way as a criticism of the current management paradigm or the work done for SEDAR 49, but rather emerge from the recognition of the potential that exists with these simulation approaches that comes out of the work conducted for SEDAR 49 and where this work could go in the future.

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

Overall, I believe the implications of the uncertainty in the technical conclusions are clearly stated. The Analytical Team clearly documented uncertainty in data sources, and, as described above, undertook many sensitivity analyses to explore the implications of decisions made when setting up the model.

TOR 5. 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 future assessments.

The Data and Assessment Workshops provided many recommendations that would be expected to increase our knowledge of the species and their assessments. The application of the DLM tools in SEDAR 49 provided a thorough review of potential management procedures and operating models for the species, but it is less clear that the method, as applied, was set up to explicitly provide catch recommendations. Personally, I believe that, if these approaches are to be used in the future (I think they should be), research towards tailoring the data-limited methods used in this assessment specifically for the provision of catch advice (as discussed under TOR's 3, 4 and 7), and specifically to the requirements for the individual species would, at least initially, lead to greater improvements in these assessments than improving data for these species. The utility of improvements to data could then be evaluated in the context of the specific operating model being used for each species.

Research towards provision of catch advice includes:

- Investigating methods to reduce uncertainty associated with life history parameter covariance,
- Developing performance metrics specific to the management requirements for each stock,
- Developing catch rules specific to the requirements for each stock,
- Separating the performance metrics and risk tolerance criteria,
- Developing optimization routines to automate the selection of targets and limits for index values, assessment frequency and associated catch recommendations, and,
- Developing methods to evaluate catch recommendations separately from the evaluation of operating models and management procedures.

The eight species assessed during SEDAR 49 are all data poor and therefore improvements to the data for these species would be expected to improve their respective assessments. In my opinion, the recommendations pertaining to improving the estimates of total removals and abundance indices are higher priorities.

With respect to removals by commercial and recreational fisheries, discard mortality and quantifying uncertainty in the discard estimates are two sources of uncertainty in the assessments. Because the removals by the fisheries have a key role in determining current abundance levels, improvements to the removal estimates would be expected to improve the catch recommendations. Spatial considerations are also important, particularly if the assumption of a single stock for each species is not valid.

Fishery-independent abundance indices do provide an indication of how a population is responding to management actions or changes in environmental conditions, and therefore provide the “real-world” test of the recommendations resulting from the simulations. Although there is the potential to further develop the DLM tool to include more than one index, I think it might be quite difficult to establish harvest catch rules (e.g., SEDAR 49 Addendum Section 3.1) with more than one index. Data that are independent of the model are necessary to evaluate the performance of the actual decisions made for a stock. In cases where more than one index is available, using indices for these evaluations might be a better use of the other indices rather than attempting to incorporate all information into a single model.

Improved information about life history parameters and fishery selectivity would be expected to improve the assessments, but this may be context specific. For example, for a bycatch species with low catch rates, uncertainty in selectivity may be relatively unimportant, whereas for a targeted species with high catch rates, the importance of uncertainty in the selectivity may be higher. An advantage of using simulation models is that the utility to improvements to the data can be explored using the model.

Many of the Assessment Workshop recommendations align with suggestions above. The ones I consider most important are:

- “Fine-tuning of the index-based and length-based methods reported herein to achieve target performance metrics”
 - This is an important recommendation particularly if it can be done via an optimization routine to automate the selection of the appropriate values.

- “Calculation and presentation of performance metrics in relation to the status quo rather than a reference method”
 - I agree with this recommendation but also encourage development of performance metrics specific to the management requirements for the stock.
- “Development of region-specific estimates of correlation coefficients for growth parameters derived from growth curves specific to the Gulf of Mexico”
 - Parameter correlation is a very important issue in these models, but I wonder if parameter correlations with steepness may play a greater role in reducing the realism of some simulated trajectories.
- “Investigation of more justifiable estimates of stock depletion such as through Productivity-Susceptibility Analysis”
 - Stock depletion is a key source of uncertainty in the assessment, but I wonder if, for bycatch species, sensitivity analyses with respect to stock depletion might be sufficient to address this uncertainty. It would be more important if a management objective was to optimise yield.
- “Estimation of current stock abundance from tagging studies (e.g. Red Drum)”
 - There is a publication (Rago 2001) that shows how one or two abundance estimates greatly enhance the utility of relative abundance indices for Atlantic Salmon. A similar analysis could be undertaken to evaluate the extent to which model performance would be improved here if abundance was known for even one or two years.
- “Evaluation of the appropriateness of target length levels which could be used in conjunction with catch and a length frequency series”
 - This is an important consideration, particularly for species for which processes other than mortality markedly influence mean length.
- “Allow for implementation error of the harvest control rule (e.g., catch recommendation overages) within the implementation model in the MSE”
 - I think this is an important recommendation, particularly for sensitivity analyses in situations where the total removals are not well understood.

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

Overall, I like the SEDAR process. I believe it provides a thorough, very well-documented and transparent mechanism for the provision of scientific information about the status of assessed fish stocks. However, because the application of DLM tools to provide catch advice as applied in SEDAR 49 is relatively new, assessing fewer species might have afforded the assessment team the opportunity to more closely focus on aspects of the assessment model, rather than updating models for several species to address issues that emerge when applying the tools to other species (I think the amount of work done by the assessment team for this assessment is quite impressive). If the DLM tools and/or related approaches are expected to be regularly used (I think the tools are a significant advancement for assessing data-limited species), then assessment of fewer species would allow the tool and its application to be more closely aligned with the assessment requirements under SEDAR, and to fine tune the application to the data and information needs for each species.

The Analytical Team also did an impressive amount of work after the Review Workshop meeting to produce the Addendum, and although I found this information very useful for understanding the behaviour of the model, I think it might also have been useful to have had the opportunity to discuss these results with the Analytical Team and Review Panel during the meeting. In SEDAR 44 (possibly a unique situation), a teleconference was held during which the Review Panel had the opportunity to discuss initial impressions of the models with the Analytical Team and to request some additional work. Although it is likely situation specific, inclusion of an initial teleconference might help both the Review Panel and Analytical Team get a head start on the work around the Review Workshop meeting.

TOR 6. 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.

I consider that the assessments undertaken during SEDAR 49 do constitute the best scientific information available for these stocks.

Relevance: The synthesis of the data and review of life history information for these stocks, as well as for stocks of the same species or closely related species, is highly relevant for the assessment of these stocks, as is the evaluation of assessment/management approaches for these species.

Inclusiveness: The assessment team did thoroughly review the available information for each stock and did consider multiple operating models. Additionally, there were opportunities for stakeholders or the public to provide input into the process.

Objectivity: The potential for subjectivity exists in virtually all scientific endeavours. In SEDAR 49, data and model decisions throughout the process were well documented, and in my opinion were made consistently and practically for the assessed species without consideration of how the decisions would influence the outcome of the assessment, and are therefore objective.

Transparency: Consistent with other SEDAR processes I have seen, I think SEDAR 49 has been highly transparent. Data, decisions, methods, and results were very well documented and the documents are in the public domain. I particularly like that as data or results are changed, corrected or updated throughout the process, the new material is produced via addendums rather than through changes to the original documents. This produces a highly transparent record of the assessment process.

Timeliness: Given the data-limited nature of these species, it is unclear whether an earlier assessment would have provided as much information. The application of the MSE approaches for the provision of catch advice for data-limited species is an emerging method, and its application here is timely.

Verification: As part of SEDAR 49, a thorough review of the all data inputs was undertaken and I consider these to be reasonably verified.

Validation: In my opinion, it is not clear how the recommended management procedures, operating models and catch advice implemented in SEDAR 49 can be validated at present. The productivity of these stocks is not known, and there are significant uncertainties in the life history model parameters used in the assessment. In the future, validation could occur via an abundance index that is not included in the model (would this be a reason, for species with only one index, to choose an operating model that does not include the index?). If either population productivity or abundance is higher or lower than expected, then the index trajectory would be expected to deviate from the expectation, although it is not clear that the cause of the deviation (abundance or productivity or both) would be identifiable without more information.

Peer review: The data limited methods used in SEDAR 49 have been peer reviewed in the scientific literature. In my opinion, the review workshop provided a thorough review of their application in SEDAR 49.

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

SEDAR 49 provided a very comprehensive presentation of the information available for the eight species included in the assessment, as well as a thorough evaluation of operating models and management procedures that could be used for these species. In my opinion, the recommendations pertaining to data collection (TOR 5) would be expected to lead to improved assessments for these species, but improvements to the modeling approaches, as discussed under TORs 3 and 4, are more important when considering the timing of the next assessment. These include:

- Investigating methods to reduce uncertainty associated with life history parameter covariance, possibly by using a different formulation for the stock-recruitment relationship or by establishing limits for metrics such as $SPR_{F=0}$, lifetime maximum reproductive rates that could provide a mechanism for filtering out unrealistic combinations of parameter values,
- Developing a wider range of performance metrics specific to each stock and separating the performance metrics and risk tolerance criteria,
- Developing methods to evaluate catch recommendations separately from the evaluation of operating models and management procedures, and
- Developing optimization routines to automate the selection of targets and limits for index values (for methods that use them), assessment frequency and associated catch recommendations.

TOR 8. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

This TOR is ongoing at the time of writing of this individual reviewer report. Writing tasks for the Peer Review Summary were assigned to the Review Panel members at the meeting and a draft Review Workshop Summary Report has been completed.

4.0. Conclusions and Recommendations in Accordance with the TOR's

SEDAR 49 was an assessment of eight data-limited species in Gulf of Mexico, including: Red Drum, Lane Snapper, Wenchman, Yellowmouth Grouper, Speckled Hind, Snowy Grouper, Almaco Jack and Lesser Amberjack. The main conclusions and recommendations from my review of this assessment are provided in this section.

TOR 1. Review any changes in data following the Data/Assessment workshop and any analyses suggested by the workshop. Summarize data as used in each assessment model. Provide justification for any deviations from Data/Assessment Workshop recommendations.

The data used during SEDAR 49 followed the recommendations of the Data and Assessment Workshops. Data used in the assessment fall into the broad categories of delineation of stocks, life history information, commercial fishery statistics (including commercial landings, discards and bycatch), recreational fishery statistics (including landings and discards), total removals, measures of fishing effort, indices of population abundance (including fishery-dependent and fishery-independent surveys), length-frequency data, age-frequency data, and other information relating to topics such as reference periods, depletion levels, species misidentification, fishery characteristics, discard mortality, and size of discards.

Despite being data-limited, the Data and Assessment workshops compiled a remarkable amount of highly relevant information for the assessment of these stocks. I think the presentation of this material was very well-organized in the workshop reports, background information and presentations, which facilitated the review.

TOR 2. 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?***
- 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?***

I generally believe that the data decisions made by the Data Workshop and Assessment Workshop are sound and robust. The uncertainties in the data were very well documented and are as expected. I believe the data are properly applied within the assessment model and that the input data series are sufficient to support the general approach used in this assessment.

Delineation of Stocks

A single stock was assumed for each species. I consider this a practical decision given the data available for the stock. However, if a single stock is assumed when more than one stock exists, stocks with lower productivity may be impacted to a higher degree, or if fishing effort is concentrated on a small number of isolated stocks, the removal rates for these isolated stocks

may not be sustainable. **A review of the distribution of fishing effort is recommended to ensure potential smaller stocks are not being inadvertently targeted.**

Life History Information

A review of the peer-reviewed literature, and published and unpublished reports, was used to identify sources of life history information; and a scoring reliability rubric was created as a semi-quantitative method to score the overall utility of the work for providing means and ranges of values for life history parameters for use in the DLM tool. I believe this approach provided a thorough and transparent process for providing life history information.

Life history parameter values are incorporated into the assessment by setting minimum and maximum values and randomly drawing values from a uniform distribution defined by these bounds. Parameter covariance is not fully addressed using this method and some combinations of parameters might be biologically unrealistic. **Where possible, choosing parameters that are not dependent on other model parameters is recommended.** For example, the steepness parameter in the stock-recruitment (SR) relationship could be replaced with the SR slope at the origin, eliminating the covariance with natural mortality rates, growth parameters, maturity parameters and length-weight conversion parameters. **Calculating metrics that are functions of several life history parameters, such as steepness, $SPR_{F=0}$, the maximum lifetime reproductive rate is also recommended to provide a mechanism for filtering out combinations of parameter values that are biologically unrealistic by using limits on these aggregate metrics.**

Commercial Fishery Statistics, Recreational Landings, and Total Removals

SEDAR 49 provided a thorough overview of the removals associated with the commercial and recreational fisheries. The reports provided good justifications for decisions about how to include these data, such as the length of the time series to use in the model. Discard mortality rates and quantifying uncertainty in the discard estimates are two sources of uncertainty in the assessments. Removals by the fisheries have a key role in determining current abundance levels. **Research towards quantifying the number of fish discarded and discard mortality rates is recommended.**

Representative Fleets and Fishing Effort

A single representative fleet and an estimate of fishing effort are required for each species in order to be able to use the DLM toolbox. The Analytical Team selected the fishery with the highest landings as the representative fleet. The number of fishing trips within a year was used as the metric for fishing effort in both the recreational and commercial fisheries. In my opinion, these are practical decisions that are well supported. Decisions about the age-at-first-capture and the age-at-full-selectivity, as well as their associated uncertainty were derived from the available data, but also included information provided by the fishermen to help ensure that the decisions were sound. **Evaluation of the sensitivity of model output with respect to very broad assumptions about selectivity is recommended to determine whether improved selectivity information would improve the assessment.**

Indices of Abundance

Abundance indices were available from both fishery-independent and fishery-dependent sources. The Data Workshop Report and background documents provided thorough descriptions of the available abundance indices, methods used for their calculation, their strengths and weaknesses, and justification for their inclusion or exclusion as data inputs for the model. No abundance indices were recommended for use in SEDAR 49 for Snowy Grouper or Speckled Hind.

Development of abundance indices for Snowy Grouper and Speckled Hind is recommended.

Reference Period and Depletion Levels

The reference period had two main functions in SEDAR 49: it was used to evaluate current practices, and also to provide a baseline status associated with the index values during that time period. Reference time periods provided in GMFMC (2011) were used for seven of the eight species. An estimate of the current level of depletion of the stock is required as an input for the DLMtool. An estimate of depletion was only available for Red Drum and estimates for the other species were derived from other, similar stocks. Sensitivity analyses for the assumed level of depletion were thoroughly done in SEDAR 49.

Although the reference periods might be appropriate as they are currently being used for management, it was not clear that they would be most appropriate for use in the model. **A review of the reference periods, particularly for species with harvest control rules that use indices, is recommended to determine if a period can be identified when the depletion level could be linked to index values (possibly pick an extreme index value?).**

Length Frequency Data

Length frequency data was obtained from both fishery-dependent and fishery-independent sources. The data and decisions about its use were well described and appear sound.

TOR 3. 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 the data-limited methods scientifically sound and robust?***
- b) Are the methods appropriate given the available data?***
- c) Are the data-limited models configured properly and used in a manner consistent with standard practices?***
- d) Are the quantitative estimates produced reliable? Does the method produce management metrics (e.g. OFL, ABC) or other indicators (e.g. trends in F or Z, probability of overfishing) that may be used to inform managers about stock trends and conditions?***

A multi-model approach was used to assess the eight stocks in SEDAR 49. The primary analytical method was the DLM Tool. Two other methods were used: a mean length-based mortality estimator assuming non-equilibrium conditions to estimate the total mortality rate, and catch curve analysis was also used to estimate total mortality. Throughout the assessment emphasis was appropriately on the DLM tool.

The DLM tool has been reviewed. It primarily uses standard equations that have been used extensively in the past, and MSE methods are established in the literature. I consider these methods to be sound.

The population simulations undertaken for SEDAR 49 are in some ways more similar to the population viability analyses used in conservation biology than they are traditional stock assessment models. These methods are also well-studied and are extensively used in planning for species-at-risk. I do believe the use of simulation models to evaluate alternative management procedures, as was done in SEDAR 49, is consistent with best practices in conservation biology. Long-term population simulations allow one to explore the logical implications of current knowledge and assumptions, and are therefore highly appropriate given the data available for the species assessed in SEDAR 49.

Code for data limited methods is rapidly evolving and some code was developed explicitly for SEDAR 49. **A review of this code to ensure it is consistent with the model equations is recommended.**

Further development of the models to reduce the effects of potentially unrealistic parameter combinations is recommended.

Further development of visual aids to guide interpretation of analytical results, such as presentation of single simulations, and measures that better should the distribution of the simulation results is recommended.

Inclusion of a quasi-extinction threshold, as is typically done in PVAs, is recommended to improve the realism of the simulations.

As applied in SEDAR 49, the data-limited methods provide an evaluation of the potential management procedures given the data available for each species. The output from these methods are options for management procedures rather than the estimates of stock size and fishing mortality rates produced by more traditional assessments. Although probability densities for catch recommendations are produced as part of the evaluation, it is not clear that they capture the full range of catch recommendations. **Fine tuning the model to ensure that index scalars, assessment frequency and other control variables in the operating model are set appropriately for the species is recommended.**

For several species, more than one management procedure met the performance criteria, while at the same time producing probability distributions for the catch recommendations associated with the procedure. The Assessment Workshop addressed this issue by producing a joint distribution that was recommended as a basis for providing management advice which may not be the best approach. **Developing a catch recommendation via specific evaluation of the management procedure to be used in the future is recommended.**

A clear interpretation of the how the catch recommendations from the DLM analyses relate to OFL, ACL, ABC, ACT or some other metric was not available at SEDAR 49. **The potential to use different percentiles from the catch recommendation distribution for different**

management metrics could be explored as a way to provide guidance on the use of the catch recommendation output.

TOR 4. 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.

Because the species assessed in SEDAR 49 are data limited, uncertainty in the data sources and assessment results is relatively high. Additionally, because the use of the DLM tool for the provision of catch advice is relatively new, further uncertainty arises from how best to use the tool for catch advice. Overall, I believe the implications of the uncertainty in the technical conclusions are clearly stated.

The workshops fully considered uncertainty in the data and operating model inputs, and incorporated uncertainty into the simulations using Monte Carlo methods. These methods are highly appropriate. Additionally, there were a large number of sensitivity analyses to evaluate the effects of the decisions were made when setting up the model. These methods capture the significant sources of uncertainty in the population, data sources, and assessment methods.

It was not clear during the review that the probability distributions for the catch recommendations were sufficiently broad. **Exploration of a broader range of catch recommendations for each operating model is recommended to ensure the uncertainty is fully quantified.**

As used for SEDAR 49, the DLM tool was set up primarily for evaluating potential management procedures and operating models, and not for evaluating whether a specific catch recommendation would meet the performance metrics. **Feeding the specific catch recommendation back into the operating model is recommended to ensure that performance metrics are met. Sensitivity analyses at this stage are also recommended.**

As implemented in SEDAR 49, the probability of meeting the performance metrics was calculated across all years and simulations independently. However, each simulation is a potential realization of future conditions that either meets management objectives (as stated as performance criteria) or does not. The probability of meeting the objective can be calculated from the proportion of simulated population trajectories that meet the criterion. **Adoption is recommended of a two two-tiered approach, including criteria that are applied to each individual simulation to determine whether it meets the metric or standard, as well as risk acceptance criteria applied across simulations based on the probability that the standard is met.**

The beauty of using these kinds of simulations is that the performance metrics can be very explicitly tied to the management objectives for the species and the relative probability of

meeting the objectives using various procedures can be used to guide future decisions. Objectives might be expected to vary depending on conditions such as whether the species is taken primarily as bycatch or is being targeted; or whether the population is in a depleted state. **On a per species basis, refinement of the questions to be addressed via the simulation model, by first stating the goals, then establishing a set of performance metrics related to these goals, and then establishing risk tolerance criteria for each of the performance metrics is recommended.**

TOR 5. 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 future assessments.

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

The application of the DLM tools in SEDAR 49 provided a thorough review of potential management procedures and operating models for the species, but it is less clear that the method, as applied, was set up to explicitly provide catch recommendations.

In my opinion, research towards tailoring the data-limited methods used in this assessment specifically for the provision of catch advice and specifically to the requirements for the individual species would be expected to lead to greater improvements in these assessments than improving data for these species, at least initially.

Research towards provision of catch advice includes:

- Investigating methods to reduce uncertainty associated with life history parameter covariance,
- Developing performance metrics specific to the management requirements for each stock,
- Developing catch rules specific to the requirements for each stock,
- Separating the performance metrics and risk tolerance criteria,
- Developing optimization routines to automate the selection of targets and limits for index values, assessment frequency and associated catch recommendations, and,
- Developing methods to evaluate catch recommendations separately from the evaluation of operating models and management procedures.

Evaluating the utility of improvements to data in the context of the specific operating model being used for each species is recommended.

I believe the SEDAR process provides a thorough, very well-documented and transparent mechanism for the provision of scientific information about the status of assessed fish stocks. However, because the application of DLM tools to provide catch advice as applied in SEDAR 49 is relatively new, assessing fewer species might have allowed for greater emphasis on the development of the model. Inclusion of an initial teleconference prior to the Review Workshop might help both the Review Panel and Analytical Team get a head start on the work around the Review Workshop meeting.

TOR 6. 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.

I consider that the assessments undertaken during SEDAR 49 do constitute the best scientific information available for these stocks.

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

In my opinion, SEDAR 49 provided a very comprehensive presentation of the information available for the eight species included in the assessment, as well as a thorough evaluation of operating models and management procedures that could be used for these species. The extent to which the research recommendations pertaining to model development are fulfilled is a main consideration when scheduling the next assessment.

TOR 8. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

This TOR is ongoing at the time of writing of this individual reviewer report.

5.0. References

- Adams, M. S., S. R. Sagarese, and A. B. Rios. 2016. Gulf of Mexico Data-Limited Species Life History Compilation. SEDAR49-DW-05, SEDAR, North Charleston, SC. 36 pp.
- Akcakaya H.R., and M.G. Raphael. 1998. Assessing human impact despite uncertainty: viability of the Northern Spotted Owl metapopulation in northwestern USA. *Biodiversity and Conservation* 7: 875-894.
- Barrowman, N. J., and R. A. Myers. 2000. Still more spawner-recruitment curves: the hockey stick and its generalizations. *Canadian Journal of Fisheries and Aquatic Sciences* 57: 665-676.
- Beissinger, S.R., and S.R. McCullough. 2002. Population viability analysis. The University of Chicago Press, Chicago, Illinois.
- Beissinger, S.R., and M.I. Westphal. 1998. On the use of demographic models of population viability in endangered species management. *Journal of Wildlife Management* 62: 821-841.
- Carruthers, T. R., and A. Hordyk. 2016. Package 'DLMtool', version 3.2. 132 pp. Available from: <https://cran.r-project.org/web/packages/DLMtool/>
- Carruthers, T. R., A. E. Punt, C. J. Walters, A. MacCall, M. K. McAllister, E. J. Dick, and J. Cope. 2014. Evaluating methods for setting catch limits in data-limited fisheries. *Fisheries Research* 153:48-68.
- Carruthers, T., L. Kell, D. Butterworth, M. Maunder, H. Geromont, C. Walters, M. McAllister, R. Hillary, P. Levontin, T. Kitakado, and C. Davies. 2015. Performance review of simple management procedures. *ICES Journal of Marine Science* 73:464-482.
- Gedamke, T., and J. M. Hoening. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosfish. *Transactions of the American Fisheries Society* 135:476-487.

- Gulf of Mexico Fishery Management Council (GMFMC). 2011. Final Generic Annual Catch Limits/Accountability Measures Amendment for the Gulf of Mexico Fishery Management Council's Red Drum, Reef Fish, Shrimp, Coral and Coral Reefs, Fishery Management Plans (Including Environmental Impact Statement, Regulatory Impact Review, Regulatory Flexibility Analysis, Fishery Impact Statement). Gulf of Mexico Fishery Management Council, Tampa, Florida. 378 pp.
- Lande, R., S. Engen, and B.-E. Saether. 2003. Stochastic Population Dynamics in Ecology and Conservation. Oxford University Press, New York.
- Lindenmayer, D.B., and H.P. Possingham. 1996. Ranking conservation and timber management options for Leadbeater's possum in southeastern Australia using population viability analysis. *Conservation Biology* 10: 235-251.
- Ludwig, D. 1999. Is it meaningful to estimate a probability of extinction? *Ecology* 80: 293-310.
- MacCall, A.D. 2002. Fishery management and stock-rebuilding prospects under conditions of low-frequency environmental variability and species interactions. *Bulletin of Marine Science* 70(2):613-628.
- McCarthy, M., M.A. Burgman, and S. Ferson. 1996. Logistic sensitivity and bounds on extinction risks. *Ecological Modelling* 86:297-303.
- McCarthy M., H.P. Possingham, J.R. Day, and A.J. Tyre. 2001. Testing the accuracy of population viability analysis. *Conservation Biology* 15: 1030-1038.
- McCarthy, M., S.J. Andelman, and H.P. Possingham. 2003. Reliability of Relative Predictions in Population Viability Analysis. *Conservation Biology* 17: 982-989.
- Millar, R.B., 2015. A better estimator of mortality rate from age-frequency data. *Canadian Journal of Fisheries and Aquatic Sciences* 72: 364-375.
- Myers, R. A., K. G. Bowen, and N. J. Barrowman. 1999. Maximum reproductive rate of fish at low population sizes. *Canadian Journal of Fisheries and Aquatic Sciences* 56: 2404-2419.
- Myers, R. A., B. R. MacKenzie, K. G. Bowen and N. J. Barrowman. 2001. What is the carrying capacity of fish in the ocean? A meta-analysis of population dynamics of North Atlantic cod. *Canadian Journal of Fisheries and Aquatic Sciences* 58: 1464-1476.
- Rago, P. J. 2001. Index measures and stock assessment in Atlantic salmon. p 137-176 *In* E. Prevost and G. Chaput [ed.], *Stock Recruitment and Reference Points, Assessment and Management of Atlantic salmon*. INRA, Paris.
- Reed, J.M., L.S. Mills, J.B. Dunning, E.S. Menges, K.S. McKelvey, R. Frye, S.R. Beissinger, M.-C. Anstett, and P. Miller. 2002. Emerging issues in Population Viability Analysis. *Conservation Biology* 16:7-19.
- SEDAR (Southeast Data Assessment and Review). 2015. SEDAR Procedural Workshop 7 - Best Practices. SEDAR. North Charleston, SC. 151 pp.
- Smith, M.W., A.Y. Then, C. Wor, G. Ralph, K.H. Pollock and J.M. Hoenig. 2012. Recommendations for catch-curve analysis. *North American Journal of Fisheries Management* 32: 956-967.
- Then, A. Y., J. M. Hoenig, N. G. Hall, and D. A. Hewitt. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. *ICES Journal of Marine Science* 72:82-92.
- Taylor, B.L. 1995. The Reliability of Using Population Viability Analysis for Risk Classification of Species. *Conservation Biology* 9 (3), 551-558.

6.0. Appendices

Appendix 1: Bibliography of Materials Provided for Review

Appendix 2: CIE Statement of Work

Appendix 3: Panel Membership

Appendix 1: Bibliography of Materials Provided for Review.

**SEDAR 49 – Gulf of Mexico Data-Limited Species
Review Workshop Document List**

Document #	Title	Authors	Date Submitted
Documents Prepared for the Data Workshop			
SEDAR49-DW-01	Shrimp Fishery Bycatch Estimates for Gulf of Mexico Data Limited Species: Wenchman and Lane Snapper, 1972-2014	Jeff Isely	6 April 2016 Updated: 20 June 2016
SEDAR49-DW-02	Catch per unit effort indices and Effort Time-series for SEDAR 49 Data Limited Species captured in the Gulf of Mexico Recreational Headboat Fishery (1986 – 2015)	Matthew S. Smith and Adyan Rios	28 April 2016
SEDAR49-DW-03	Timeseries of effort and nominal abundance indices derived from the Gulf of Mexico recreational private and charter fishery for the species included in the SEDAR 49 data limited stock assessment	Matt Smith	Not Received
SEDAR49-DW-04	Review of bycatch in the Gulf menhaden fishery with implications for the stock assessment of red drum	Skyler R. Sagarese, Matthew A. Nuttall, Joseph E. Serafy and Elizabeth Scott-Denton	27 April 2016
SEDAR49-DW-05	Gulf of Mexico Data-Limited Species Life History Compilation	Molly S. Adams, Skyler R. Sagarese, and Adyan B. Rios	18 April 2016
SEDAR49-DW-06	Lane snapper <i>Lutjanus synagris</i> Findings from the NMFS Panama City Laboratory Trap & Camera Fishery-Independent Survey 2004-2014	D.A. DeVries, C.L. Gardner, P. Raley, and K. Overly	22 April 2016
SEDAR49-DW-07	The Red Drum (<i>Sciaenops ocellatus</i>) spawning population in the eastern Gulf of Mexico: composition, site fidelity, and size	Susan Lowerre-Barbieri, Mike Tringali, Joel Bickford, Sarah Burnsed, and Mike Murphy	20 April 2016
SEDAR49-DW-08	Summary of length data and length frequency distributions for eight data limited species collected in the Gulf of Mexico from 1981 to 2015	Ching-Ping Chih	27 April 2016

Appendix 1: Bibliography of Materials Provided for Review.

SEDAR49-DW-09	SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Almaco Jack	Matthew D. Campbell, Kevin R. Rademacher, Paul Felts, Brandi Noble, Joseph Salisbury, John Moser, Ryan Caillouet	29 April 2016
SEDAR49-DW-10	SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Lane Snapper	Matthew D. Campbell, Kevin R. Rademacher, Paul Felts, Brandi Noble, Joseph Salisbury, John Moser, Ryan Caillouet	29 April 2016
SEDAR49-DW-11	SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Lesser Amberjack	Matthew D. Campbell, Kevin R. Rademacher, Paul Felts, Brandi Noble, Joseph Salisbury, John Moser, Ryan Caillouet	29 April 2016
SEDAR49-DW-12	SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Snowy Grouper	Matthew D. Campbell, Kevin R. Rademacher, Paul Felts, Brandi Noble, Joseph Salisbury, John Moser, Ryan Caillouet	29 April 2016
SEDAR49-DW-13	SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Speckled Hind	Matthew D. Campbell, Kevin R. Rademacher, Paul Felts, Brandi Noble, Joseph Salisbury, John Moser, Ryan Caillouet	29 April 2016
SEDAR49-DW-14	Size Composition of Eight SEDAR49 Data Limited Species by Sector and Gear	J.J. Isely, M.W. Smith and C-P Chih	3 May 2016
SEDAR49-DW-15	Almaco jack <i>Seriola rivoliana</i> Findings from the NMFS Panama City Laboratory Trap & Camera	D.A. DeVries, C.L. Gardner, P. Raley, and K. Overly	29 April 2016

Appendix 1: Bibliography of Materials Provided for Review.

	Fishery-Independent Survey 2004-2014		
SEDAR49-DW-16	Current Status of Adult Red Drum (<i>Sciaenops ocellatus</i>) in the North Central Gulf of Mexico: An Update of Abundance, Age Composition, and Mortality Estimates	Crystal L. Hightower, J. Marcus Drymon, and Sean P. Powers	2 May 2016 Updated: 8 May 2016
SEDAR49-DW-17	Lane Snapper Abundance Indices from SEAMAP Groundfish Surveys in the Northern Gulf of Mexico	Adam G. Pollack, David S. Hanisko and G. Walter Ingram, Jr.	2 May 2016 Updated: 11 May 2016
SEDAR49-DW-18	Wenchman Abundance Indices from MSLABS Small Pelagics Surveys in the Northern Gulf of Mexico	Adam G. Pollack, David S. Hanisko and G. Walter Ingram, Jr.	2 May 2016 Updated: 11 May 2016
SEDAR49-DW-19	Wenchman Abundance Indices from SEAMAP Groundfish Surveys in the Northern Gulf of Mexico	Adam G. Pollack, David S. Hanisko and G. Walter Ingram, Jr.	2 May 2016 Updated: 11 May 2016
SEDAR49-DW-20	SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Wenchman	Matthew D. Campbell, Kevin R. Rademacher, Paul Felts, Brandi Noble, Joseph Salisbury, John Moser, Ryan Caillouet	4 May 2016
SEDAR49-DW-21	SEAMAP Reef Fish Video Survey: Relative Indices of Abundance of Yellowmouth grouper	Matthew D. Campbell, Kevin R. Rademacher, Paul Felts, Brandi Noble, Joseph Salisbury, John Moser, Ryan Caillouet	4 May 2016
SEDAR49-DW-22	Summary of length and weight data for seven data limited species collected during NMFS and SEAMAP fishery-independent surveys in the Gulf of Mexico	David S. Hanisko and Adam Pollack	20 May 2016
Documents Prepared for the Assessment Process			
SEDAR49-Assessment Workshop-01	Review of Operating Model Parameters for SEDAR 49: Almaco Jack	Skyler R. Sagarese, J. Jeffery Isely, and Matthew W. Smith	12 July 2016
SEDAR49-	Review of Operating Model	Skyler R. Sagarese,	12 July 2016

Appendix 1: Bibliography of Materials Provided for Review.

Assessment Workshop-02	Parameters for SEDAR 49: Lane Snapper	J. Jeffery Isely, and Matthew W. Smith	Updated: 12 August 2016
SEDAR49-Assessment Workshop-03	Review of Operating Model Parameters for SEDAR 49: Lesser Amberjack	Skyler R. Sagarese, J. Jeffery Isely, and Matthew W. Smith	12 July 2016
SEDAR49-Assessment Workshop-04	Review of Operating Model Parameters for SEDAR 49: Red Drum	Skyler R. Sagarese, J. Jeffery Isely, and Matthew W. Smith	12 July 2016 Updated: 12 August 2016
SEDAR49-Assessment Workshop-05	Review of Operating Model Parameters for SEDAR 49: Wenchman	Skyler R. Sagarese, J. Jeffery Isely, and Matthew W. Smith	12 July 2016 Updated: 12 August 2016
SEDAR49-Assessment Workshop-06	Review of Operating Model Parameters for SEDAR 49: Yellowmouth Grouper	Skyler R. Sagarese, J. Jeffery Isely, and Matthew W. Smith	12 July 2016 Updated: 12 August 2016
SEDAR49-Assessment Workshop-07	Synthesis of Literature on Von Bertalanffy Growth Parameter Correlations	Nancie Cummings, Skyler Sagarese and Bill Harford	29 July 2016
SEDAR49-Assessment Workshop-08	Review of Operating Model Parameters for SEDAR 49: Speckled Hind	Skyler R. Sagarese, J. Jeffery Isely, and Matthew W. Smith	12 August 2016
SEDAR49-Assessment Workshop-09	Review of Operating Model Parameters for SEDAR 49: Snowy Grouper	Skyler R. Sagarese, J. Jeffery Isely, and Matthew W. Smith	12 August 2016
SEDAR49-Assessment Workshop-10	Technical description of operating models in data-limited methods toolkit (DLMtool)	William J. Harford, Skyler R. Sagarese, J. Jeffery Isely, and Matthew W. Smith	30 August 2016
Documents Prepared for the Review Workshop			
SEDAR49-RW-01	Revised Results for the Generic Implementation of Itarget0 and Ltarget0 for Lane Snapper, Wenchman, Lesser Amberjack, and Almaco Jack	Skyler R. Sagarese, J. Jeffery Isely, and Matthew W. Smith	21 October 2016
Final Stock Assessment Reports			
SEDAR49-SAR1	Gulf of Mexico Data-limited Species	SEDAR 49 Panels	
Reference Documents			
SEDAR49-RD01	Spatial and size distribution of red drum caught and released in Tampa Bay, Florida, and factors associated with the post-release hooking mortality	Kerry E. Flaherty, Brent L. Winner, Julie L. Vecchio, and Theodore S. Switzer	
SEDAR49-RD02	Evaluating the current status of red	Crystal LouAllen Hightower	

Appendix 1: Bibliography of Materials Provided for Review.

	drum (<i>Sciaenops ocellatus</i>) in offshore waters of the North Central Gulf of Mexico: age and growth, abundance, and mercury concentration	
SEDAR49-RD03	DLMtool: Data-Limited Methods Toolkit (v3.2)	Tom Carruthers and Adrian Hordyk
SEDAR49-RD04	Evaluating methods for setting catch limits in data-limited fisheries	Thomas R. Carruthers, André E. Punt, Carl J. Walters, Alec MacCall, Murdoch K. McAllister, Edward J. Dick, Jason Cope
SEDAR49-RD05	Evaluating methods for setting catch limits in data-limited fisheries: Supplemental Appendix A	Thomas R. Carruthers, André E. Punt, Carl J. Walters, Alec MacCall, Murdoch K. McAllister, Edward J. Dick, Jason Cope
SEDAR49-RD06	Performance review of simple management procedures	Thomas R. Carruthers, Laurence T. Kell, Doug D. S. Butterworth, Mark N. Maunder, Helena F. Geromont, Carl Walters, Murdoch K. McAllister, Richard Hillary, Polina Levontin, Toshihide Kitakado, and Campbell R. Davies
SEDAR49-RD07	Performance review of simple management procedures: Supplemental Appendix A	Thomas R. Carruthers, Laurence T. Kell, Doug D. S. Butterworth, Mark N. Maunder, Helena F. Geromont, Carl Walters, Murdoch K. McAllister, Richard Hillary, Polina Levontin, Toshihide Kitakado, and Campbell R. Davies
SEDAR49-RD08	Generic management procedures for data-poor fisheries: forecasting with few data	H. F. Geromont and D. S. Butterworth

Appendix 2: CIE Statement of Work.

Statement of Work

External Independent Peer Review by the Center for Independent Experts

SEDAR 49 Gulf of Mexico Data-Limited Species 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 49 will be a compilation of data, an assessment of the stock, and CIE assessment review conducted for Gulf of Mexico Data-Limited Species. 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 49 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 17 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 **Miami, FL from November 1-3, 2016**.

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

Appendix 2: CIE Statement of Work.

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.

CIE reviewers shall conduct an impartial and independent peer review of the assessment in accordance with the SoW and ToRs herein.

Appendix 2: CIE Statement of Work.

A description of the SEDAR Review process can be found in the SEDAR Policies and Procedures document:

http://sedarweb.org/docs/page/SEDARPoliciesandProcedures_Oct15_FINAL_update.pdf

The CIE reviewers may contribute to a Summary Report of the Review Workshop produced by the Workshop Panel.

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 in Miami, Florida from November 1-3, 2016, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
- 3) No later than December 2, 2016, each CIE reviewer shall submit an independent peer review report addressed to the “Center for Independent Experts,” and sent to Dr. Manoj Shivilani, CIE Lead Coordinator, via email to mshivilani@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**.

Appendix 2: CIE Statement of Work.

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

September 27, 2016	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
October 18, 2016	NMFS Project Contact sends the CIE Reviewers the pre-review documents
November 1-3, 2016	Each reviewer participates and conducts an independent peer review during the panel review meeting
December 2, 2016	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
December 16, 2016	CIE submits CIE independent peer review reports to the COTR
December 23, 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 be 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.

Appendix 2: CIE Statement of Work.

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
Northern Taiga Ventures, 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

Appendix 2: CIE Statement of Work.

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.

Appendix 2: CIE Statement of Work.

Annex 2: Terms of Reference for the Peer Review

SEDAR 49 Gulf of Mexico Data-limited Species Review Workshop

1. Review any changes in data following the Data/Assessment workshop and any analyses suggested by the workshop. Summarize data as used in each assessment model. Provide justification for any deviations from Data/Assessment Workshop recommendations.
2. 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 Assessment Workshop sound and robust?
 - b) Are data uncertainties acknowledged, reported, and within normal or expected levels?
 - c) Are data applied properly within the assessment model?
 - d) Are input data series reliable and sufficient to support the assessment approach and findings?
3. 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 the data-limited methods scientifically sound and robust?
 - b) Are the methods appropriate given the available data?
 - c) Are the data-limited models configured properly and used in a manner consistent with standard practices?
 - d) Are the quantitative estimates produced reliable? Does the method produce management metrics (e.g. OFL, ABC) or other indicators (e.g. trends in F or Z, probability of overfishing) that may be used to inform managers about stock trends and conditions?
4. 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.
5. 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 future assessments.
 - Provide recommendations on possible ways to improve the SEDAR process.
6. 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.
7. Provide guidance on key improvements in data or modeling approaches that should be considered when scheduling the next assessment.
8. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference.

Appendix 2: CIE Statement of Work.

Annex 3: Tentative Agenda

SEDAR 49 Gulf of Mexico Data-limited Species Review Workshop
November 1-3, 2016
Miami, Florida

Tuesday

9:00 a.m. Introductions and Opening Remarks <i>- Agenda Review, TOR, Task Assignments</i>	Coordinator
9:30 a.m. – 11:30 a.m. Assessment Presentations <i>- Assessment Data & Methods</i> <i>- Identify additional analyses, sensitivities, corrections</i>	Analytic Team
11:30 a.m. – 1:00 p.m. Lunch Break	
1:00 p.m. – 6:00 p.m. Assessment Presentations (continued) <i>- Assessment Data & Methods</i> <i>- Identify additional analyses, sensitivities, corrections</i>	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 <i>- Assessment Data & Methods</i> <i>- Identify additional analyses, sensitivities, corrections</i>	Chair
11:30 a.m. – 1:00 p.m. Lunch Break	
1:00 p.m. – 6:00 p.m. Panel Discussion/Panel Work Session <i>- Continue deliberations</i> <i>- Review additional analyses</i> <i>- Recommendations and comments</i>	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 <i>- Final sensitivities reviewed.</i> <i>- Projections reviewed.</i>	Chair
11:30 a.m. – 1:00 p.m. Lunch Break	
1:00 p.m. – 5:30 p.m. Panel Discussion or Work Session <i>- Review Reports</i>	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.

Appendix 3: Panel Membership

Review Panel Membership

Luiz Barbieri, Chair	Gulf SSC
Panayiota Apostolaki	CIE Reviewer
Yong Chen	CIE Reviewer
Jamie Gibson.....	CIE Reviewer
Kai Lorenzen.....	Gulf SSC
Joe Powers	Gulf SSC