Center for Independent Experts (CIE) Independent Peer Review of

Gulf of Mexico Data-Limited Species Assessment Review SEDAR 49

Prepared by:
Panayiota Apostolaki
January 2017

Assignment undertaken for:
Center for Independent Experts, CIE, US

Table of Content

Executive Summary

The Gulf of Mexico data-limited Species Review Workshop took place in Miami between November 1st and November 3rd, 2016. The focus of the review was the results of the assessments carried out for eight species; Red Drum, Lane Snapper, Wenchman, Yellowmouth Grouper, Speckled Hind, Snowy Grouper, Almaco Jack, and Lesser Amberjack.

All eight species were considered to be data-poor so, running a conventional stock assessment would be of little value. For this reason, Management Strategy Evaluation (MSE) was used to test the performance of multiple data-limited models for stock assessment and development of management advice for the 8 species of interest. The DLMTool provided the software the assessment team used to conduct the MSE analysis.

The assessment team also considered a mean-length estimator model as well as catch curve analysis as an alternative way to calculate metrics that could support management decisions. However, the former produced results that were not considered reliable while the latter could only be applied to one species (red drum) since the necessary information was not available for the rest of the species.

Three external reviewers from the Center of Independent Experts (CIE) attended the Review meeting and considered the assessment process and its outcomes. The assessment team provided a very useful description of the work and also undertook additional analysis in response to questions from the Review Panel. The review aimed to ascertain whether the assessment could underpin management decision-making. To do so, a number of factors were considered including the robustness of the outcomes, the appropriateness of the approach, and relevance of calculated metrics.

The results showed that index-based and mean-length-based management procedures (MPs) could be a good choice for the eight species considered and, for the majority of scenarios tested, those MPs performed better than the MP currently in place.

The MSE analysis was an appropriate approach that allowed inference about effective management plans to be made despite the fact that a conventional stock assessment could not be done. However, this approach did highlight challenges not only in terms of knowledge gaps but also with the interpretation and use of outcomes.

This is because the analysis did not produce metrics and indicators that managers are used to seeing (e.g. the conventional stock assessment will provide absolute numbers for OFL, Fmsy, etc.). So, this could pose a challenge for managers since they might need to adjust the decision-making process to facilitate the uptake of the new outcomes, but also for scientists in terms of finding the most effective way to present the new metrics. The analysis also highlighted knowledge gaps to which the methods considered are more sensitive and that can guide discussions about future research and monitoring and prioritisation of tasks.

Background

The 2016 Gulf of Mexico Data-limited Species Assessment (SEDAR 49) Review Workshop focused on eight species found in the Gulf of Mexico and are exploited either through targeted fishing or caught as by-catch. The species considered are Red Drum, Lane Snapper, Wenchman, Yellowmouth Grouper, Speckled Hind, Snowy Grouper, Almaco Jack, and Lesser Amberjack. 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. SEDAR 49 is comprised of a compilation of data, an assessment of the stock, and a CIE assessment review.

For all eight species there is limited information about their life history and exploitation patterns. A stock assessment had been done for one of the species (red drum), but even for that species the analysis of available information indicated that it is not a data-rich species.

Key parameters characterised by paucity of information included natural mortality, stock structure, and growth. With regards to fisheries, some catch data were available; the longest catch time series was for red drum (1981 to present) while the shortest was for wenchman (1997 to present). Recreational catches were also constructed using information from sampling programmes. The longest time series recommended for inclusion in the assessment was for red drum (1981 to present) and the shortest for speckled hind and wenchman (1997 to present). Estimates of discards were also constructed. CPUE series were developed using these data and/or data from fishery-independent surveys. One CPUE series was chosen for each species (using the most representative fishery for each stock) and was included in the assessment.

Data-limited techniques were employed to explore the effectiveness of different management approaches for each of the species. Specifically, the assessment employed a management strategy evaluation approach using the DLMTool software. This approach was followed because running a conventional stock assessment would be of little value due to the data-poor nature of the eight species. The assessment team also considered a mean-length estimator model, as well as a catch curve analysis, as alternative ways to calculate metrics that could support management decisions. However, the former produced results that were not considered reliable while the latter could only be applied to one species (red drum), since the necessary information was not available for the rest of the species.

The data limited toolbox (DLMTool) was developed by Carruthers and co-workers (Carruthers *et al.* 2014; Carruthers & Hordyk 2015) and is an R package that contains the standard MSE components of an operating model and data-limited management procedures. The MSE approach included an age-structured model which was used as the operating model, and a suite of management procedures (MPs) that could be used for data limited species. The assessment team elected to test a sub-group of those MPs that were deemed more appropriate for the data available for the eight species.

The outcomes of the assessment focused on the performance of the different MPs and sensitivity analysis was used to test the effects of uncertainty on model outcomes and identify parameters and assumptions that had the biggest impact on

findings. The results of the assessment are presented in section III of the final SEDAR 49 report.

Three CIE reviewers were commissioned to conduct an impartial and independent peer review of the 2016 Gulf of Mexico Data-limited Species Assessment (SEDAR 49) in accordance with the SoW and stock assessment ToRs listed in Appendix 2. Each CIE reviewer is also required to produce an independent peer review report and may also assist the Chair of the Panel review meeting with contributions to the Summary Report, based on the terms of reference of the Review.

This document provides my review of the work covered in the Gulf of Mexico Data-limited Species Assessment and supporting material that underpinned that work. Further details on the reviewer's role and the review request of the Center for Independent Experts are presented below and in Appendix 2.

Description of the Reviewer's Role in the Review Activities

I was contracted to:

1) Conduct necessary pre-review preparations, including reviewing necessary background information and reports provided by the NMFS Project Contact in advance of the peer review.
2) Actively participate during the panel review meeting as a member of the review panel and complete the independent peer review according to required format and content as described in Annex 1 (Appendix 2, Annex1). Complete the independent peer review addressing each ToR as described in Annex 2 (Appendix 2, Annex 2).
3) Assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review.
4) No later than 2nd December 2016, submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Dr. Manoj Shivlani, CIE Lead Coordinator and Dr. David Sampson, CIE Regional Coordinator.

Summary of findings

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 stock assessment team conducted additional analysis during the Review Workshop and afterwards to respond to questions from the Panel. Part of that analysis involved using different parameters from those adopted at the Data/Assessment workshop. The additional runs and analyses are described in the document entitled *"Addendum: SEDAR 49 Gulf of Mexico Data-limited Species"* which forms part of the SEDAR 49 report. The main changes considered aimed to:

- test the sensitivity of results to the values of key parameters including parameter beta, the duration of the assessment interval, and natural mortality;
- explore the impact of assumptions that underpinned the assessment such as the choice of reference period for the harvest control rules and the way in which performance metrics were calculated;
- assess the impact of inter-annual variability in the growth model and greater uncertainty in the observation error for the index of abundance and
- consider additional diagnostics.

In terms of deviations from the Data/Assessment workshops, runs under the first three categories listed above diverged from the recommendations of the Data/Assessment Workshops as describe below.

- The change in the value of parameter beta aimed at simplifying the original formula by removing the effect of hyper-stability or hyper-depletion from the performance of different management approaches. This is useful when interpreting the results as it makes it easier to spot trends in results that depart from the expected behaviour.
- The assessment interval was set equal to 10 years; that period of time does not reflect well the biology of some of the species assessed and it was considered that waiting for 10 years to adjust the management plan runs the risk of missing important changes in the stock status. A smaller assessment interval was chosen (3 years) to create a scenario that allowed for more frequent evaluation of stock status.
- Model runs using alternative values for M were conducted to test the performance of the management approaches. The reason for considering different values for M was because the Panel felt that basing the calculation of M on a single approach (Hoenig's formula) did not represent common practice.
- The value of steepness has not been estimated directly for any of the species, but is a key parameter affecting the productivity of a stock. By deviating from the original parameterisation and fixing the steepness to its lower value, we can do a simple test to check what management approaches remain viable even if a very conservative assumption about stock productivity (i.e. low productivity) is used.

- The addition of inter-annual variability in the results of the growth model aimed to increase realism in the simulation by allowing for plasticity in growth values and reflected research findings on the topic.

All the additional assumptions and changes considered added value to the assessment and helped clarify questions and lead to better understanding of the potential of the approach employed and robustness of associated scientific advice. The findings from the additional runs together with the original analysis are discussed under some of the other ToR below.

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 DW and AW sound and robust?
b. Are data uncertainties acknowledged, reported, and within normal or expected levels?
c. Are data applied properly within the assessment model?
d. Are input data series reliable and sufficient to support the assessment approach and findings?

Data used for the assessment include stock identification and life history information, fishery statistics, CPUE series, and length frequency data. Meta-analysis and use of values from similar or temperate species were also part of the data gathering exercise to fill gaps in knowledge. Assumptions about stock depletion were also used as input data to the assessment. The information available varied depending on the species but remained limited across the group of species.

Growth parameters were available for six of the eight species and a von Bertalanffy growth model was used for all of them. That model is generally a reasonable choice. However, it is important to note that it was not considered to be the most appropriate for red drum as it does not describe the gender-changing characteristic of that species, but alternative models that provided a better fit could not be simulated by the DLMTool. Therefore, it is recommended that future assessments use models that better describe the growth pattern of this species.

Similarly, although a von Bertalanffy model was created for speckled hind, there is significant uncertainty about the maximum age for this stock which could be underestimated in the parametrisation adopted. For Wenchman, data presented indicate that there is further analysis that could be done to produce more robust estimates. Therefore, further work to better estimate the age span of speckled hind and Wenchman (including validation of annual deposition) is recommended as it will affect the growth parameters used and resilience of the species to overexploitation.

Catch at length data presented also raised concerns as they included fish that were considerably bigger than those predicted by the corresponding von Bertalanffy growth model. Although, this might not be as important for this assessment given that the aim was to simulate a stock that would broadly resemble each of the eight considered, it is recommended that further work is done to improve growth estimates. This will be important in the future and will make conventional stock assessment models more realistic.

For the stock structure, information was available for only two of the species (red drum and lane snapper) but the assumption of a single stock was used for all eight

species. Given the paucity of information to suggest otherwise, this assumption is the simplest one, and essentially, reflects current knowledge (or lack of it).

The decision to use a single equation (the Hoenig equation) to calculate mortality, although it was taken after evaluation of robustness of different methodologies, is questionable especially because maximum age for some of the eight species could not be defined or there were concerns about the values used. Thus, natural mortality could not be estimated for some stocks. Consideration of other models could have also provided an alternative way to define the uncertainty around point estimates of M. Therefore, it would be useful to consider more than one approach to calculate M, but it is acknowledged that the assessment team conducted sensitivity analyses that considered model performance across a range of M values to account for uncertainty.

The decision to use meta-analysis to find estimates of steepness when those were not available is a reasonable choice given the lack of information about this parameter. However, it is noted that the value of steepness has not been estimated directly for any of the stocks; so, this is an important source of uncertainty. The assessment used sensitivity analysis to mitigate for that but given the importance of this parameter it is recommended that consideration is given to alternative ways to calculate steepness or reparametrize the stock recruitment function to use other parameters for which plausible range of values might be easier to define.

Catch series were available for all eight species but the reliability of the data varies among species, and that became more apparent for species that are caught as by-catch (non-target species) or are misreported (e.g. Yellowmouth grouper). It is not straightforward to get the full picture about catches since some of the data are confidential, but that aside, the steps followed to compile the data and identify issues with incomplete reporting or misreporting are appropriate.

Estimates of by-catch in the shrimp fishery were calculated using observers' data for lane snapper and wenchman, and in the bottom longline for snowy grouper and speckled hind. On bycatch, the assumption that lane snapper has low discard mortality needs to be confirmed, as it currently relies on anecdotal information (expert testimony). The same holds for lesser amberjack.

The methodology followed to calculate recreational catches and discards is appropriate and makes use of available data while recognising important weaknesses in the data collection program. It is noted that discard mortality values rely mainly on anecdotal information while there is little knowledge about the length frequency of discards. However, for some of the species considered (e.g., lane snapper) recreational catches are the main component of fishing pressure. Therefore, increasing the reliability of this part of the catch data is needed.

Uncertainty has been assigned to catch and by-catch data, and that goes some way to address concerns about robustness of the data series, but it cannot reflect year/period specific uncertainty as the model in the DLMTool accepts a single uncertainty estimate for the entire time series for each species.

Length frequency data were available for five of the eight species, namely, red drum, lane snapper, wenchman, snowy grouper, and speckled hind, and the information was used to build selectivity curves by identifying the length at full gear selectivity. This is a simple way to develop selectivity curves and makes the most of the information available.

Effort was calculated for the representative fishery for each stock either from commercial fleet records such as logbooks or using estimates of angler trips for the recreational fisheries to support development of CPUE indices. Indices of abundance were also produced using fishery-independent survey data. One CPUE series was chosen for each species (corresponding to the representative fishery), and although the steps followed to decide on the preferred CPUE series are appropriate, the small CV characterising some of the CPUE series raises concerns about the robustness / representativeness of the data series.

Some of the management plans required selecting a time period from the years covered in each CPUE series. That time period was used as a reference period to parametrise index-based and other management approaches. For red drum, such a period was not specified during the DW and therefore, the five last years in the time series were used as the reference period. This is a reasonable choice. However, evidence is not available to suggest that this period represents a phase during which the population has recovered. So, it is suggested that parameter values calculated to represent this reference period (e.g. CPUEref) are used as a lower boundary in the management plans (or upper boundary if it is catches) rather than a target. That will be at least for the short term, and until additional information becomes available to provide more clarity about the status of the stock.

Overall, there are a number of gaps/weaknesses in the data used to describe the life history and exploitation of the species but they do reflect and capture current knowledge well. The MSE approach used here does not require well defined estimates of the input parameters and therefore, concerns about some of the adopted values are less material for this phase of the assessment. However, this will not be the case once/if absolute values of stock size, optimum exploitation yield, etc. are required to guide management, and therefore, research to fill some of the major gaps still remains a priority. It is recognised though that it is unlikely to fill all the important gaps (e.g., past information on recreational catches for lane snapper).

Despite serious data limitations, the data used and the process by which the input data were developed represent a sound and appropriate way forward. Some changes were needed and the assessment team responded efficiently to recommendations from the reviewers to improve data inputs and processes used. Significant data gaps remain, but the data used provided an acceptable basis for the MSE approach.

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?

The Assessment workshop explored different data-limited assessment models and management procedures using a Data Limited Methods toolbox (DLMtool) to evaluate the potential of different Management Procedures (MP) for each of the eight species.

The DLMtool contains the standard MSE components of an operating model and data-limited management procedures. The MSE facilitates simulation testing of uncertainties and biases in the data and life history parameters/assumptions. This toolkit and the management procedures are freely available and the different components of the model have been peer reviewed through the journal publications process, for example Carruthers *et al.* (2014); Geromont and Butterworth (2014) (see Reference documents).

A subset of the DLMtool Management Procedures was tested and their choice reflected the data available for each of the eight species. The Management Procedures all require a different mix of information. For example, the catch only methods require a catch series and information on depletion, whereas the index-based method mostly needs a recent index of abundance (e.g., CPUE).

Overall, the methods used represent reasonable choices given the paucity of data that limits the spectrum of quantitative approaches that could be used. The adoption of a simulation evaluation approach in the context of the DLMTool allows for a relatively abstract and high level consideration of management procedures which reflects the knowledge gaps. This approach adds value to the assessment since it gives an indication of the procedures that might be fit for the nature of species and fisheries studied. However, the DLMTool and MSE approach are not meant to replace standard stock-assessments and their use does not mean that less effort should made to improve knowledge about the species life-history and exploitation.

The adopted approach does pose certain challenges through both in terms of the way it is implemented/designed, and also its application to develop scientific advice and guide management.

For the former, the main issue is that significant effort has been put into populating the DLMTool and collecting the data that will fit that approach. Because of that, it seems that less thinking went into considering the nature of exploitation and how the eight species fit in the bigger picture/fishery to define the analysis that would be best suited to assess the impact of exploitation on those species. As such, it should be highlighted that the methods used were designed for target species, probably overexploited, and slow-growing. The stocks considered in this assessment include by-catch or non-target species with shorter life spans (only some of the eight species), but could be highly susceptible to fishing. Furthermore, it is not clear whether the main aim of the fishery (or management policy) is to achieve catches at MSY or reduce inter-annual variability (two of the metrics considered).

Interpretation and use of the results of the assessment to guide management could also pose a challenge because this requires a different management paradigm as the tested methodology does not produce the metrics that are calculated in a conventional stock assessment (e.g., Bmsy).

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

The DLMToolkit is the main package used in this analysis and the management approaches chosen were those that made use of the information available such as

relative abundance information in the form of a CPUE index or mean length information to provide a signal about the magnitude of future exploitation. The main equations in the population dynamics model are standard formulae that have been used extensively in the past and are scientifically sound. The management procedures (MPs) considered have also been used in other studies and peer-reviewed as part of previous work (e.g., Geromont and Butterworth, 2014) so, the general concept is sound. However, those are empirical MPs and their original parameterisation, as used in the DLMtool, has been adopted to support management of severely depleted stocks. In addition to the DLMtool approaches, catch curve analysis was also employed to calculate total mortality.

In principle, there is good understanding of the scientific basis and use of the operating model and MPs in the DLMtool, but there are still concerns about the implementation of the approach. Specifically:

- Some parts of the code, especially those involving internal boundaries and checks that are hardwired into the code, are not fully explored and could lead to diversions from the main formulae and influence the results. An example is the automatic adjustment of the fishing mortality to avoid extinction, hence producing a more optimistic picture of the fishing impacts. Those internal adjustments need to be checked and documented in detail to ensure that the performance of tested MPs is not artificially enhanced.
- The approach is still under development and requires a very good understanding of the underlying concepts and their translation into source code, and that takes a lot of time so, it is not a quick shortcut to assessing data-poor species.
- The scalars of the formulae used to describe the MPs tested in this assessment have been chosen to provide a resilient approach to help overexploited species to recover and achieve MSY. Therefore, there is no evidence that the same values for the scalars represent the best option for the type of species assessed in this exercise.

b) Are the methods appropriate given the available data?

The methods proposed aim to address paucity of biological and other information in data limited species. All the species considered here could be assigned to that category so, the use of DLMtool is generally appropriate. The volume and extent of data for red drum was relatively greater than for the other species, and that warrants further consideration to decide whether this stock can be treated as a data-moderate one. This does not render the Toolkit inappropriate, but suggests that consideration of additional quantitative approaches could be of value to identify those that are more suitable.

In terms of whether the current methods provide the most appropriate or flexible mechanism for capturing the characteristics of species and fisheries considered here, it should be noted that the methods and parameterisation of the models has been designed for target species and at least half of the species considered in this set of assessments were not targeted species. So, the type of information available or of use could differ from that for targeted species. For example, effort patterns or future catch quota characterising the target species in the relevant fishery in which the study species are caught is an alternative source of information that could guide projections. However, it is not clear whether the current configuration of the model

can make use of such information. Therefore, it is recommended that an alternative approach also be considered in the future. The alternative configuration will use information about the status of the target species of the fishery in which this species is caught to develop plausible effort scenarios and identify the management approach that performs best.

Length frequency data were also used with an extended version of the Beverton-Holt length-based mortality estimator, but the estimates of total mortality were not considered reliable due to problems with conversion or representativeness of the inputs of the population described. Therefore, this method is not examined further in this report.

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

The values of the model parameters reflect the recommendations of the data workshop and in that sense they are properly configured, but given concerns about parameter values selected by the data workshop, there are recommendations for further work to address them. Those include:

- The choice of L_{inf} is not supported by catch at length data that for most of the species considered appear to include considerably higher values for fish length.

- Similarly, the CV for the growth parameters are unrealistically small so, this part of the model configuration needs to be revisited.

- All simulations adopted the reference periods indicated by the management authorities and used them to determine changes in future catches. However, there is very little information about the state of nature that reference period represents and no clear justification for its selection. In conventional assessments, the reference period is set at a much earlier time and is assumed to either reflect the state of the population that led to optimum production or, in some cases, the state of the population at almost unexploited conditions. The interpretation of model predictions will be affected by those assumptions and therefore, the choice of the reference period needs to be substantiated and an explanation provided of what state of nature it is supposed to represent.

- Work is still needed to justify the choice and parameterisation of the HCRs. Linked to this, Figure 2 in Addendum suggests that recommended catches under Itarget and w=0 are non-zero. However, the corresponding equation suggest that catches will be 0 if w=0. This should be revisited.
- The timeframes for the simulations does not reflect the dynamics of some of the stocks (e.g., 40 years for a species that lives for five years). It is recommended that the simulation time be calculated as a function of generation time or a similar constant to reflect the biology of the assessed stock.

Additional analyses requested by the reviewers highlighted that the HCR tended to favour higher catches in the short term and postponed recovery for later in the projection period. Results from individual runs, but also over a large number of runs show a behaviour that is unexpected. Some runs seem to delay the recovery

(reduction in F) of the stock for several years (e.g., for lane snapper it is 10 years) or do not reduce effort at all (Fig 5.3, 5.5 in Addendum). The expectation is that the HCR will reduce the pressure when the stock is overexploited so, these simulations exhibit the opposite behaviour.

This raises questions about the accuracy and appropriateness of the MPs to produce a pattern of exploitation control that will be in line with requirements under the Magnuson-Stevens Act and associated guidelines. Even when a great number of simulations are used to estimate each of the parameters the behaviour produced is unexpected, specifically, the model results show that the HCR will continue to cut effort well after the biomass has increased to sizes well above Bmsy. Again, this does not fit the behaviour that the mathematical formula for the HCR describes (e.g. Figures 4.3, 5.2 in the addendum).

From the relevant documents and discussions during the review workshop, it transpired that more data than those used for red drum existed. This suggests that the model for red drum does not reflect the best available knowledge. It is understandable that given the big number of species (eight species) that had to be assessed, compromises in the data compilation, and hence model configuration were inevitable but that weakened the value of the analysis. It is recommended that future assessments allow enough time to identify and compile all available data to improve the model configuration.

Furthermore, the assessment team was asked to produce joint probability density functions (pdf) for catch recommendations by combining catches that come from different MPs. There is no clear justification for this decision and it results in catches that have not been tested in the MSE exercise. As such, this configuration is questionable and does not represent standard practices. So, it is recommended that catches are produced for each MP separately, unless further explanation and analysis is provided to justify the appropriateness and representativeness of this approach.

The combination of I_0scalar =1. And I_{target} scalar =1 is not a permitted configuration for Itarget0 method (New Table 5.3A) so, this should be excluded from the calculations (and also check why the model does produce results for that combination).

d) Are the quantitative estimates produced reliable? Does the method produce management metrics (e.g. MSY, ABC, ACL) 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?

Within the context of data limited approaches, the results of the assessment can guide management decisions on future allowable catches. However, the outcome of this analysis does not correspond to the traditional metrics produced in data rich assessments (e.g., MSY or OFL).

The estimates produced in this assessment mainly concern metrics that describe the performance of alternative management procedures. That is judged based on indicators such as the probability of the population and yield to be above a pre-specified reference point (MSY), probability of not overfishing, and probability that the biomass will not fall below a pre-specified limit for each of the MPs considered.

The assessment could be configured to provide catch estimates and more conventional indicators, but it is not clear how they relate to ABC, OFL, etc. As such,

the metrics are useful, but the way in which they will be incorporated into the decision-making process has not been clarified yet and it is expected to require an adjustment in the current procedure for setting catch quotas. Nevertheless, this is a valuable approach that can guide decisions and help avoid overexploitation while the knowledge is built to develop a robust assessment. It can do so by providing signals about stock status and exploitation levels in the absence of absolute estimates about stock size and exploitation.

In terms of informing management decisions, the choice of performance indicators does not reflect well the fact that most of the species are by-catch/no-target species. For example, it is questionable whether achieving maximum yield is a realistic or relevant objective in these fisheries. Hence, more work is needed to define how the relevant indicators are expected to inform management decisions and whether all indicators that have been calculated should be given the same weight when one decides on the best MP to use.

Therefore, it is recommended that performance metrics and additional criteria are revisited and possibly adjusted to reflect the fact that these stocks are discard species, and because of that, certain objectives such as avoiding overexploitation could be more important or relevant than achieving maximum yield.

The evaluation outcomes were tested under a range of scenarios and uncertainty levels and the main conclusions were not affected. That provides some assurance about the robustness of the estimates and the reliability of the outcomes of the management evaluation in terms of the MPs that are more appropriate for the assessed stocks. However, the influence of the constraints of the model (see previous section about hardwired checks in the source code) on probability density functions reduces the reliability of the results.

TOR 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.

Uncertainty in input data and assumptions used is considerable. This includes key biological parameters such as growth parameters and steepness, and fisheries data including catches and abundance indices. To reflect that, the DW has calculated CVs and recommended sensitivity analyses. Where the calculated CVs were deemed to be small or CVs had not been calculated, a CV based on expert judgement was chosen. The model also requires information about plausible ranges of stock depletion for each of the species and the assessment used relevant information from similar species to fill that knowledge gap as such estimates were not available for the eight stocks. Reliability scores were also assigned to each data input to guide decisions on methods available in the DLMTool that could be used to provide management advice.

Monte Carlo simulations were used to propagate the uncertainty and characterize it in the model outputs; one thousand runs were done using values for the model parameters that were randomly drawn from pdfs reflecting the range of values and uncertainty that the DW or AW had assigned to each parameter. Model stability and convergence was assessed using convergence plots for the performance metrics.

The assessment examined the sensitivity of the methods' performance to assumptions about stock depletion in the operating model to show how sensitive they were to the values of that parameter and help understand the implications if the level of depletion is misspecified. It also did runs with different assumptions about total removals during the reference period to assess the sensitivity of catch recommendations to the magnitude of total removals. The configuration of the model also accounted for correlation between the growth parameters to make the search across the space of plausible value combinations more efficient and realistic.

Overall, this represents considerable effort to reflect uncertainty in the data and the use of MSE is appropriate as it allows for uncertainties to be captured, but also facilitates evaluation of alternative management option(s) despite, or in the light of, uncertainties in data inputs and assumptions. This provides an informative picture of the implications of uncertainty in the outputs of the analysis.

Shortcomings associated with the input data and methodology have been covered above and are also relevant for this section; some key points relating to uncertainty are covered below:

- There is a single CV representing uncertainty in CPUE; that is a limitation imposed by the configuration of the models in the DLMTool, but a single CV is unlikely to represents reality as information presented suggests the quality of data has improved in recent years, at least for some of the species.

- The CV assigned to some of the input parameters was unrealistically small (e.g., growth parameters). However, additional runs were conducted during the review workshop to add variability in the growth model and test its effect on model outcomes, so this point was partially addressed during the assessment meeting.

- The use of biological information from temperate species to define plausible parameter values for the species considered in this assessment might have introduced bias in the range of plausible values.

- The choice of steepness represents one of the key areas in which knowledge is lacking, so that adds to the uncertainty around model outputs. Admittedly, it is very difficult to get good estimates of this parameter and the assessment had used a distribution of values to capture uncertainty.

TOR 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.

a) Research and monitoring recommendations

The list of recommendations for research and improvement of data quality both from the DW and AW is considerable and reflects the gaps in knowledge but also the gap in research and data collection programs that focus on the eight species assessed. These recommendations are reasonable and aim to produce data that would support a conventional stock assessment. As mentioned already, the use of a MSE approach

does not eliminate the need to assess the status of the stock, and therefore, the research recommendations are justifiable.

I have included below some of the additional work that responds to the findings of the analysis and will improve the reliability of future assessments

In terms of priorities, the assessment results indicated that HCRs that rely on indices (CPUE or length-based) show more promise in supporting management advice for the species assessed. Therefore, better data collection and fishing records to improve collection of effort, catches and size distribution information is recommended. However, as some species might be misidentified (e.g., lesser amberjack) future work should aim to assess the magnitude of this problem; such knowledge will improve future data, but could also be used to adjust past catch data allocated to each stock.

Linked to this, it is recommended that the SEAMAP indices for snowy grouper and speckled hind are revisited to ascertain whether further work could be done to make them suitable for use in the assessment of the two species. Also, a scheme to produce CPUE indices for the recreational fishery for red drum could help provide a more complete picture of the exploitation.

It is clear that information about discard mortality is lacking, and for some species, the impact of different fisheries is not well mapped (e.g., red drum, lane snapper), so some impact might be overlooked. Additional work is needed to quantify that either through new schemes or by extending existing relevant data collection programmes. At the same time, it is not clear whether personal records of fishermen could provide relevant information, but it is one approach that merits consideration.

Mapping the impact of fisheries on other than targeted species could also help develop a more ecosystem-focused approach or identify species from the complex of species that a fishery impacts that could be used as indicators of the health of all species in the complex.

b) Recommendations on possible ways to improve the SEDAR process

The SEDAR process has been running for some time now, so it is well streamlined. However, this assessment was different as it involved species that are not the primary target in a fishery or the target at all. During the meeting, it was clear that the thinking and approach needed to be adjusted to accommodate those new aspects of the SEDAR process.

The approach followed seemed to assume that all eight species could be clustered together in a single assessment which usually is used to assess two or three species. However, given these are species not previously assessed and also do not fall in the conventional group of species, they require more time to first discuss the best approach to assess them and then implement that decision. The fact that the DLMTool was used and it offered off-the-shelf models did not reduce the time required to produce a reliable and meaningful assessment. Actually, as it transpired, the use of the DLMTool approach required a lot of work and time. Therefore, my recommendation will be to avoid including so many species in a single SEDAR, especially when species that have not been assessed before are involved.

Because of the nature of the species and fisheries the assessment covered, it would have been good to know why each species was chosen for the assessment. It was

not clear from the material reviewed what criteria were used to identify these species, what specific challenges or problems the selection group considered, and what the priorities of the assessment were. It would be useful if future SEDAR reviews include such information, especially for species that have not been assessed before.

For data-poor species, I would also recommend that the review process includes an extra day dedicated to input data which will cover all the assumptions and decisions made to fill knowledge gaps and produce the data for the stock assessment.

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.

The assessment offers a very useful picture of the situation with the eight species considered and it represents the most advanced analysis conducted for those species so far. It has used the most appropriate modelling approach to address data limitations and a well-thought process for defining the values of the parameters needed for the analysis. Therefore, it does offer the best scientific information.

Even though there are considerable gaps in knowledge, the choice of the MSE approach provides an objective way to inform management discussions and identify effective management strategies. It is also timely, as it considered management approaches that are more elaborate than those currently in use. Thus, it can guide work on defining an effective management framework and focus discussions on additional data and research that needs to be prioritized.

The assessment has made use of relevant and up-to-date knowledge on assessment and management evaluation of data-poor species; they have employed an MSE approach, which as a methodology is well documented and tested and the software used is freely available, so everyone could use it to replicate the analysis adding to the transparency of the work. The use of meta-analysis to fill knowledge gaps is appropriate and reflects efforts to include all relevant information.

The process has been informed by both scientific literature/research as well as expert knowledge, and the DW and AW made an effort to incorporate both quantitative and qualitative data and make use of fishermen's experience, increasing the inclusiveness of the assessment. The process also considered both fishery dependent and fishery independent information and the latter offers an alternative verification avenue. Unfortunately, some of the fishery independent surveys could not be used and that is an area where improvements could be made as it could increase the robustness of the analysis.

Steps have also been taken to verify and validate results by considering different methods for calculating key indicators (e.g., fishing mortality); these were the models in the DLMTool, a length-based model, and catch curve analysis. However, some additional work is still needed to fully understand the behavior of the model, especially some of the hard-wired assumptions and checks to ensure that the model reproduces the right behavior.

Although the overall approach is sound and offers an effective way to make progress, there are two aspects of this work that can improve; these are the parameterization of the HCRs and the performance metrics. Both are currently based on targeted, overexploited species, which does not necessarily produce the best fit for the species

considered in this assessment. This also relates to the objectives for the fisheries assessed. Again, it is not clear what the aim is and how different objectives rate in terms of priority (e.g., achieve MSY, avoid bottlenecks, ensuring that those stocks are not going extinct).

So, although the technical approach in this assessment is sound, conceptually, I believe there is a step before that to provide clarity on the priorities underpinning the stock assessment of these species. This will help build a much more tailored assessment and will increase its relevance.

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

Additional work to improve data quality has been discussed under ToR 5. A few points about modelling approaches and the assessment methodology used are included below.

As mentioned in earlier sections of this report, additional work to refine the performance indicators and better define the projection period used to test the performance of HCRs is needed. For the latter, it is important to align the timeframe in which the HCR needs to deliver positive change with any regulatory or other constraints that an MP will need to meet in reality. As the analysis showed, 40 years might be a very long period to use to test whether a HCR is working especially for short lived species. Similarly, it was not clear which indicators and metrics were the most appropriate to use to assess the performance of management strategies. This needs to be clarified before the next assessment.

In addition, the behavior of some of the HCRs needs further examination as they do not seem to create patterns of stock recovery or catch trajectories that are in line with the general expectations, such as that if a stock is overexploited the HCR will lead to catch reduction. So, first the formulae themselves need to be explored further to ensure that they are fit for purpose and are tuned to the fisheries assessed (since they were developed for different stocks/fisheries), and second the source code needs to be checked to ensure that the formulae have been programmed correctly.

Linked to the above, I agree with the recommendation of the AW for further work to evaluate the appropriateness of target catch or index levels which could be used in conjunction with catch and index time series. The reference period used for the HCRs needs to be discussed further, especially in conjunction with the parameterisation of the HCRs to ensure that there is consistency, for example, that the abundance index from a reference period chosen to represent pristine conditions is not used as a target in the HCR.

The fact that the DLMTool accepts only one CPUE series is a limitation. That might be particularly relevant in the case of Speckled Hind and snowy grouper since the index of advantage for each of them does not appear to be reliable when used as a single series given changes in fleet behaviour that took place during the period it covers. Therefore, it is suggested that this aspect of the DLMTool be extended to accept more than one index of abundance.

Conclusions/Recommendations

The SEDAR 49 Review Workshop (RW) took place on 1-3 November 2016. The meeting reviewed the results of applying data-poor models to provide scientific advice for eight species selected by the Gulf of Mexico Fishery Management Council. The stock assessment team also performed additional analysis during and after the Review Workshop to respond to questions from the reviewers.

The focus of the assessment was on testing various data-limited methods for providing management advice even though traditional management metrics (e.g., ABC) could not be produced. This is to overcome limitations in the data and high uncertainty that means that a conventional stock assessment could not be done.

The assessments showed that index-based methods and mean-length methods were viable approaches for providing management advice and, in the majority of the cases considered, they performed better than the management plan currently in place. A length based model which was also considered did not work well with the input data provided and was not used for performance evaluations.

Overall, the assessment approach chosen was appropriate and made good use of the information available. Specifically:

- The data used and the process followed to define the values of the input parameters represent a sound way forward.
- The data used reflect and capture current knowledge well but also highlight the significant number of gaps in understanding of the life history of the species and the fisheries that exploit them.
- The findings of the assessment provide a useful insight into management plans that could be effective and can inform management discussions. However, further tuning of the MPs could improve their representativeness and realism.
- Although considerable uncertainty still remains, model results indicated that the MPs currently used to manage the exploitation of the eight species might not be the most appropriate ones and other MPs could perform better.

The assessment and Data Panel produced a long list of recommendations for further work that reflects the number of gaps in knowledge for all species considered. All research recommendations included in the assessment report are valid and will help fill those gaps.

Recommendations made under each of the ToR (Summary of Findings) in this review are also listed below:

Recommendation 1: Research to fill some of the major knowledge gaps about life history and exploitation is a priority, but it is recognised that it is unlikely to fill all the important gaps (e.g., past information on recreational catches for Lane snapper).

Recommendation 2: It is recommended that future assessments use models that include more than one parameterisation for the growth function to better capture the growth pattern of species like red drum, for which the von Bertalanffy model is not considered to be the most appropriate.

Recommendation 3: Further work is needed to better estimate the age span of the species assessed, especially speckled hind and Wenchman (including validation of annual deposition), since length frequency data appear to support greater maximum length than that currently used in the growth function.

Recommendation 4: It would be useful to consider more than one approach to calculate M as that would provide a better picture of the range of values that might be plausible. It is acknowledged though that the assessment team conducted sensitivity analyses that considered model performance across a range of M values to account for uncertainty.

Recommendation 5: Future work needs to consider alternative ways to calculate steepness or reparametrize the stock recruitment function to use other parameters for which plausible ranges of species-specific values might be easier to define.

Recommendation 6: The assumption that lane snapper has low discard mortality needs to be confirmed as it currently relies on anecdotal information (expert testimony). The same holds for lesser amberjack.

Recommendation 7: Given that for some of the species considered (e.g., lane snapper) recreational catches are the main component of fishing pressure, it is important that more work be done to consider ways in which the reliability of this part of catch data could be increased.

Recommendation 8: At least in the short term and until more clarity about the stock status can be provided, it is suggested that the values (e.g., CPUEref) calculated for the reference period (last 5 years) and used to parameterise MPs for red drum be treated as lower bounds rather than targets.

Recommendation 9: Internal checks and adjustments hardwired in the source code of the DLMTool need to be explored further and documented in detail to ensure that the performance of tested MPs is not artificially enhanced.

Recommendation 10: The volume and extent of data for red drum is slightly greater than for the other species and that warrants further consideration to decide whether this stock can be treated as a data-moderate one.

Recommendation 11: An alternative model configuration is suggested for species that are by-catch, which will use information about the status of the target species of the fishery in which this species is caught to develop plausible future effort scenarios and used them to identify management approaches that performs better.

Recommendation 12: The growth parameters need to be revisited; this is because information presented suggest that there are additional data that can be analysed to inform the estimation of those parameters, but also to address concerns about L_{inf} as, at least for some of the species, it seems to be smaller than observed values for fish length.

Recommendation 13: More generally, it is recommended that future assessments allow enough time to identify and compile all available data to strengthen the model configurations as it seems that more data than those used in the assessment were available but could not be analysed in time.

Recommendation 14: There is very little information about the state of nature that the reference period used in the MPs represents and no clear justification for the choice of that reference period. The choice of the reference period needs to be substantiated and an explanation provided regarding what state of nature it is supposed to represent.

Recommendation 15: The exploitation control patterns exhibited by some of the MPs considered is not in line with conventional behavior that, for example, would lead to catches being reduced if a stock is overexploited. So, further work to understand the behavior of those MPs is suggested.

Recommendation 16: The timeframes for the simulations does not reflect the dynamics of some of the stocks (e.g., 40 years for a species that lives for five years). It is recommended that simulation time be calculated as a function of generation time or a similar constant to reflect the biology of the assessed stock.

Recommendation 17: Furthermore, it is important to align the timeframe in which the HCR needs to deliver positive change with any regulatory or other constraints that an MP will need to meet in reality.

Recommendation 18: Further explanation is needed to justify the development of joint pdfs for catch recommendations by combining catches that come from different MPs, as such a decision results in catches that have not been tested in the MSE exercise.

Recommendation 19: It was not clear which indicators and metrics were the most appropriate to use to assess the performance of MPs. It is recommended that performance metrics and additional criteria are revisited and possibly adjusted to reflect the fact that some of these stocks are discard species.

Recommendation 20: There is a single CV representing uncertainty in CPUE; it is unlikely that such a configuration represents reality as information presented suggests the quality of data has improved in recent years, at least for some of the species. So, it is recommended that future assessments allow for year/period specific CVs.

Recommendation 21: It is suggested that the DLMTool be extended to accept more than one index of abundance.

Recommendation 22: Better data collection and fishing records to improve collection of effort, catches and size distribution information are recommended. However, as some species might be misidentified (e.g., lesser amberjack) future work should also aim to assess the magnitude of this problem; such knowledge will improve future data quality and could be used to adjust past catch data allocated to each stock.

Recommendation 23: Some of the fishery independent surveys could not be used and that is an area where improvements could be made. For example, it is recommended that the SEAMAP indices for snowy grouper and speckled Hind be revisited to ascertain whether further work could be done to make them suitable for use in the assessment of the two species.

Recommendation 24: Additional work is needed to improve information about discard mortality and about the impact of different fisheries, either through new

schemes or by extending existing relevant data collection programmes. It is not clear whether personal records of fishermen could provide relevant information, but it is one approach that merits consideration.

Recommendation 25: It is recommended that a smaller number of species be included in a single SEDAR, especially when species that have not been assessed before are involved.

Recommendation 26: It would be useful if future SEDAR reviews included information to explain why each species was chosen for the assessment; especially for species that have not been assessed before. This could cover things such as criteria used to identify the species and the priorities/objectives of the assessment.

Recommendation 27: For data-poor species, I would also recommend that the review process includes an extra day dedicated to input data which will cover all the assumptions and decisions made to fill knowledge gaps and produce the data for the stock assessment.

Recommendation 28: It is not clear what the aim of the assessment was and how different objectives rate in terms of priority (e.g., achieve MSY, avoid bottlenecks, ensuring that those stocks are not going extinct). Additional work to provide clarity on the priorities underpinning the stock assessment will be of value and will help build a much more tailored assessment and increase its relevance.

Recommendation 29: I agree with the recommendation of the AW for further work to evaluate the appropriateness of target catch or index levels which could be used in conjunction with catch and index time series.

Appendix 1: Bibliography

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 indies 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 *Lutjanus synagris* 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 (*Sciaenops ocellatus*) 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-	Summary of length data	Ching-Ping Chih	27 April

DW-08	and length frequency distributions for eight data limited species collected in the Gulf of Mexico from 1981 to 2015		2016
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 *Seriola rivoliana* 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	29 April 2016
SEDAR49-DW-16	Current Status of Adult Red Drum (*Sciaenops ocellatus*) 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-	Summary of length and	David S.	20 May

DW-22	weight data for seven data limited species collected during NMFS and SEAMAP fishery-independent surveys in the Gulf of Mexico	Hanisko and Adam Pollack	2016
Documents Prepared for the Assessment Process			
SEDAR49-AW-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-AW-02	Review of Operating Model Parameters for SEDAR 49: Lane Snapper	Skyler R. Sagarese, J. Jeffery Isely, and Matthew W. Smith	12 July 2016 Updated: 12 August 2016
SEDAR49-AW-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-AW-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-AW-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-AW-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-AW-07	Synthesis of Literature on Von Bertalanffy Growth Parameter Correlations	Nancie Cummings, Skyler Sagarese and Bill Harford	29 July 2016
SEDAR49-AW-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-AW-09	Review of Operating Model Parameters for SEDAR 49: Snowy Grouper	Skyler R. Sagarese, J. Jeffery Isely, and	12 August 2016

		Matthew W. Smith	
SEDAR49-AW-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 drum (*Sciaenops ocellatus*) in offshore waters of the North Central Gulf of Mexico: age and growth, abundance, and mercury concentration	Crystal LouAllen Hightower	
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	Thomas R. Carruthers, Laurence T. Kell, Doug D. S.	

	procedures	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. Statement of Work for Dr Panagiota Apostolaki

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.

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.

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**.

5) 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.

6) 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**).
7) 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 Shivlani, CIE Lead Coordinator, via email to mshivlani@ntvifederal.com, and Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.
8)

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 completed with the format and content in accordance with **Annex 1**,
(2) The CIE report shall address each ToR as specified in **Annex 2**,
(3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

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

Support Personnel:

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

Manoj Shivlani, CIE Lead Coordinator
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

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

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

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

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

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

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

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

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

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference 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 AW sound and robust?
   b) Are data uncertainties acknowledged, reported, and within normal or expected levels?
   c) Are data applied properly within the assessment model?
   d) Are input data series reliable and sufficient to support the assessment approach and findings?
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.

Annex 3: Agenda

SEDAR 49 Gulf of Mexico Data-limited Species Review Workshop

November 1-3, 2016
Sonesta Coconut Grove Hotel, Miami, Florida

Tuesday

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

- Agenda Review, TOR, Task Assignments

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

- Assessment Data & Methods

- Identify additional analyses, sensitivities, corrections

11:30 a.m. – 1:00 p.m. Lunch Break

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

- Assessment Data & Methods

- Identify additional analyses, sensitivities, corrections

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

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

Wednesday

8:00 a.m. – 11:30 a.m. Panel Discussion Chair

- Assessment Data & Methods

- Identify additional analyses, sensitivities, corrections

11:30 a.m. – 1:00 p.m. Lunch Break

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

- Continue deliberations

- Review additional analyses

- Recommendations and comments

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

Thursday

8:00 a.m. – 11:30 a.m. Panel Discussion Chair

- Final sensitivities reviewed.

- Projections reviewed. Chair

9:00 a.m. Introductions and Opening Remarks

Coordinator

- *Agenda Review, TOR, Task Assignments*

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

- *Assessment Data & Methods*

- *Identify additional analyses, sensitivities, corrections*

11:30 a.m. – 1:00 p.m. Lunch Break

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

- *Assessment Data & Methods*

- *Identify additional analyses, sensitivities, corrections*

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

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

Wednesday

8:00 a.m. – 11:30 a.m. Panel Discussion Chair

- *Assessment Data & Methods*

- *Identify additional analyses, sensitivities, corrections*

11:30 a.m. – 1:00 p.m. Lunch Break

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

- *Continue deliberations*

- *Review additional analyses*

- *Recommendations and comments*

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

Thursday

8:00 a.m. – 11:30 a.m. Panel Discussion Chair

- *Final sensitivities reviewed.*

- *Projections reviewed.* Chair

11:30 a.m. – 1:00 p.m. Lunch Break

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

- *Review Reports*

5:30 p.m. – 6:00 p.m. Public comment Chair

6:00 p.m. ADJOURN

Thursday Goals: Complete assessment work and discussions, final results available.

Draft Reports reviewed.

Appendix 3: List of participants

Workshop Panel

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

Analytic Representation

Skyler Sagarese ..... SEFSC, Miami
Jeff Isely ..... SEFSC, Miami
Shannon Cass-Calay ..... SEFSC, Miami

Appointed Observers

Ben Blount ..... Gulf SSC
Claudia Friess ..... Gulf Appointee

Attendees

Shanae Allen ..... FWRI
Jay Grove ..... FWC
Bill Harford ..... Univ. of Miami
Matthew Johnson ..... SEFSC
Mike Larkin ..... SERO
Michekke Masi ..... FWRI
Kevin McCarthy ..... SEFSC
Michael Schirripa ..... SEFSC
Matthew Smith ..... SEFSC
Beth Wrege ..... SEFSC

Staff

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