Center for Independent Experts
Consultant Report on:

SEDAR 22 Stock Assessment Review
Gulf of Mexico Yellowedge Grouper and
Gulf of Mexico Golden Tilefish
February 14 - 17, 2011

Paul A. Medley
Sunny View
Jack Hole
Alne, YO61 1RT
UK
paulahmedley@yahoo.co.uk

Contents

Executive Summary

SEDAR 22 covered Gulf of Mexico yellowedge grouper (*Epinephelus flavolimbatus*) and Gulf of Mexico golden tilefish (*Lopholatilus chamaeleonticeps*) stock assessments. As well as this report, a SEDAR 22 Review Panel Report contains a consensus review of the assessments and the scientific advice. This report covers the activities and views of the reviewer, Paul Medley. It does not conflict with any findings in the review panel report. The main findings, recommendations and conclusions are:

- The data and assessments reported by the review panel form a good basis for scientific advice. Uncertainties associated with the assessment are dealt with appropriately and taken into account in the advice. The results represent the best scientific advice available for these stocks.
- The data used were appropriate, and suitable for use in stock assessment. The main weakness identified was the treatment of the abundance indices. Possible improvements in the standardisation of the available abundance index data were identified.
- The main assessment software, Stock Synthesis 3 (SS3), is appropriate for the type of data available and should enable the scientists to set up assessments and obtain credible and consistent results. The results from the SS3 assessments were valid and provided a good basis for management advice.
- Additional work was still being conducted before the completion of this report. The consensus Review Workshop Final Report will have a summary of final assessment results. Based on the available assessment results, the yellowedge grouper biomass and fishing mortality is very close to the current benchmark ($SPR_{30\%}$), whereas golden tilefish biomass and fishing mortality is above the current benchmark. It both cases, catches should be adjusted downward, possibly more sharply in the golden tilefish case, to achieve the management objectives.
- The abundance index standardisation could be improved, with the objective of reducing the estimated index observation errors and ensuring that the scale of the observation errors is consistent with the age/length composition observation errors.
- The SS3 reference age parameter for the natural mortality-at-length should be fixed, based on the relevant ages provided for the natural mortality estimate. The way this might be done is described.
- The large number of sensitivity runs to scope the uncertainty in the assessment presents a problem for combining this uncertainty into a single form for advice. I suggest reducing sensitivities as much as possible, though developing priors for key parameters, such as natural mortality, rather than using fixed values over a range. This should make better use of the stochastic simulation facility (MCMC) in SS3.

Background

SEDAR 22 consisted of assembling the relevant fisheries data sets for Gulf of Mexico Yellowedge Grouper and Golden Tilefish, two benchmark stock assessments, and an assessment review. These stocks are within the jurisdiction of the Gulf of Mexico Fishery Management Council and the states of Texas, Louisiana, Mississippi, Alabama, and Florida.

This report concerns the final review workshop of the SEDAR process. A Data Workshop (DW) followed by an Assessment Workshop (AW) had already been conducted. The DW develops and approves the data and some model parameters for use in the stock assessment. The AW develops and approves the stock assessment model configuration as well as deciding which data are to be used. Reports are produced by both these workshops for inclusion in the review.

The review workshop (RW) provides an independent peer review of SEDAR stock assessments and is responsible for ensuring that the best possible assessment is provided through the SEDAR process. The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise for this review through the Center for Independent Experts (CIE: www.ciereviews.org). Three CIE reviewers were selected to conduct the independent peer review of NMFS science in compliance the Terms of Reference (see Appendix II). Each of three CIE reviewers was contracted to deliver an independent peer review report to be approved by the CIE Steering Committee.

Description of Review Activities

The independent peer review covers the data, assessment models, and results previously developed for and by the data and assessment workshops. The SEDAR documents include working papers prepared for each workshop, supporting reference documents, and the SEDAR Stock Assessment Reports.

Before the formal review took place, the background material and reports were provided by the NMFS Project Contact in advance. The information provided included the Data Workshop and Assessment Workshop reports (see Appendix I).

The review panel, of which I was a member, met at the Embassy Suites Hotel in Tampa, FL, from 1:00 p.m. Monday, February 14, 2011 through 1:00 p.m. Thursday, February 17, 2011. The review evaluated assumptions, identified errors and improvements to the stock assessment which could be implemented within a reasonable time frame. The review panel has completed a consensus report on their findings for each stock, subject to final analyses being undertaken by the assessment team. This review report will contain a summary of the stock assessment results, including suitability of the data and model and the results and the uncertainty associated with those results. The external reviewers were Robin Cook (UK), Paul Medley (UK), and Henrik Sparholt (UK, CIE). Stephen Szedlmayer (USA) was also on the panel as a local representative and the panel was chaired by Douglas Gregory (USA).

The review meeting consisted of a series of presentations on the stock assessments which had been conducted. Linda Lombardi presented the methods and results of the stochastic stock reduction analysis (SRA). Brian Linton presented the methods and

results of the Stock Synthesis 3 (SS3) assessment and other analyses of tilefish. John Walter presented the methods and results of the SS3 assessment and other analyses of yellowedge grouper. After the presentations, the reviewers were given an opportunity to ask questions, request further results from the assessment, as well as additional stock assessment runs or sensitivity analyses. The review panel requested one additional sensitivity analysis for each assessment. No significant errors were identified, but a number of improvements, in my opinion, were suggested.

This report is not the consensus review report, but an independent peer review report which addresses the same Terms of Reference in the consensus report. This report does not disagree with any finding in the consensus review report, but contains additional information and opinions which are not necessarily a consensus view. Some additional work was conducted after the review panel met to help elucidate problems and issues which arose during the meeting. The intent of this report is to be constructive in terms of recommendations for the future direction of research and development of the stock assessment and the SEDAR process.

Summary of Findings

1 Evaluate the adequacy, appropriateness, and application of data used in the assessment.

All data are appropriate, and suitable for use in stock assessment. Data collection has been standardised with recorded catches, a longline survey and sampling of length and age conditional on length. This basic data collection scheme should be suitable for all species caught by the fishery from these waters. The main weakness is the limited sampling, and therefore data, available for these stocks.

Life history of these species is not well understood, particularly for tilefish. As information improves, it is likely that the model will need to be changed to reflect improved understanding.

The catch data available for both tilefish and yellowedge had problems. Early historical catches were not recorded to species level, so various assumptions had to be made to divide these catches among species. This was particularly a problem for yellowedge grouper. However, dividing these early catches was done in a reasonable way, and it is difficult to see how this might be improved. Discards and recreational catches while uncertain, were clearly small in these fishery, and therefore not a significant source of uncertainty for the assessment.

Length and age compositions are limited in quantity, but are a significant source of information on these stocks since 2000. Given the model is trying to estimate stock size and catch-at-age by sex and area, data are particularly sparse. Ageing errors are high. Beyond expansion of the data collection and sampling programmes, it appears that few improvements can be made.

The main correctable problems with the data were the abundance indices. These were based on bottom-set longline for surveys and the commercial fishery. I agree that not using the SEAMAP trawl index was appropriate because catches of the target stocks were too low to produce a useful index from this source. The quality and quantity of the data varies between sources, with the highest quality and lowest quantity from the

bottom-set longline survey. While the quality of all series seems as good as possible, I believe that treatment of the data could be improved.

The approach to standardise the indices was to use generalized linear models (GLMs) to remove variation not associated with changes in abundance. This was appropriate, but the GLMs could be improved. In particular, the variance on each index was relatively large, which reduces the information that the abundance indices bring to the assessment. One of the aims of standardisation should be to try to reduce this error.

The standardisation used did not change the general trends in the nominal CPUE index, so although the performance was not as good as it could be, the indices were not incorrect. The parameter correlation matrix, which was unavailable, would need to be inspected to ensure heavily correlated parameters are not present or used together for calculating the standardised index.

Additional covariates for use in the standardisation were suggested by the Data Workshop, and might include habitat indicators and physical oceanography (e.g. water temperature). These would be unlikely to change trends in the series, but could reduce measurement error on the abundance indices.

If possible standardisation should avoid the delta-lognormal approach, which was used. This is not parsimonious, and while possibly justified for trawl-based indices, it seems inappropriate for longline. Using the logarithm of the data is likely to lead to over-weighting smaller catches, and makes dealing with zeros more complex than it needs be. A better approach to modelling the likelihood would be to use a function based on either the binomial or Poisson probability densities. For a constant soak time and even density, the total catch in numbers of fish will form a binomial distribution where the number of trials is the number of hooks. Taking account of a multi-species catch, a multinomial likelihood can be used, which can be dealt with as a Poisson. Over-dispersed versions of these likelihoods (such as beta-binomial, negative binomial or over-dispersed Poisson) would be more likely to work, but the fact that the likelihood allows zeroes as observations suggests that they are not only more appropriate, but will be more parsimonious.

One of the reasons delta-lognormal is used is to cope with two modes (one at zero and another greater than zero). However, if the high frequency of zeroes cannot be explained by covariates within the GLM, it could suggest the selection of targeted sets is flawed. A way around this which should be explored is to analyse all catches from all species simultaneously, which could adjust for targeting using all species catch with other covariates. This might make more sense as the availability of hooks to catch fish depends on catch of all fish whether of interest for this assessment or not, as well as bait loss.

The effects on the catch rate are likely to be multiplicative, and therefore a log link function is appropriate. A good general form for the GLM linear predictor of a longline set would be:

$$lp = a_1 \ln(Hooks) + a_2 SoakTime \cdots$$

In this case, a_1 could be either estimated close to 1.0 or forced to be 1.0, making catch proportional to the number of hooks in a multiplicative model and a_2 would likely be negative, implying a diminishing chance of empty hooks over longer soak times.

Estimating a value for a_1 slightly less than 1.0 allows for local depletion and hook interference, and greater than 1.0 allows for the targeting of higher fish density (this should only apply for the commercial fishery). Values different from 1.0 are dangerous, however, as the model may be attempting to account for changes in CPUE due to abundance changes. If required, terms a_1 and a_2 should be fitted whether statistically significant or not. After these terms, the remaining linear predictor would represent the catch-per-hook and be directly interpretable as catchability q. Any fit should also be presented with residuals plotted against expected values and other standard diagnostics.

Another advantage of this GLM is that it becomes relatively simple to develop a multinomial model for multispecies (see McCullagh and Nelder 1989 Pg 211). This model could estimate the catch conditional on the total catch (all species) and the total catch conditional on the number of hooks.

I would not recommend using generalized additive models (GAM), except for more complex tasks for spatial modelling such as building density maps. The additional smoothing option does not, in my experience, help in standardising abundance indices.

If year-interaction terms are used in a model, having them as random effects (as in this assessment) is probably the best option. However, the additional random effects assumption does not eliminate the problem of the potential bias in the estimated trend if the year main effects terms are being used as the index. However, there was no evidence of this bias in the indices generated for this assessment.

2 Evaluate the adequacy, appropriateness, and application of methods used to assess the stock.

The main assessment software, Stock Synthesis 3 (SS3), is appropriate for the type of data available and should enable the scientists to set up assessments and obtain credible and consistent results.

The stochastic stock reduction analysis (SRA) was carried out and compared to the full SS3 assessment. Apart from building confidence in SS3 outputs, SRA was not particularly useful to the review. This is not to say it might be used as by the assessment workshop to help develop the SS3 model and improve confidence in its results. However, the population model was very similar and much of the same data was used, including data which was considered an important source of uncertainty (i.e. total catches). Other assessments which only used one or two sources of data might have been more useful, such as biomass dynamics models (total catch and abundance indices), length-converted catch curves, and catch curves (which were carried out).

The results from the SS3 assessments were valid and provided a good basis for management advice. The lack of clear cohorts in the age structure means that the model will be smoothing over the age structure rather than picking out the individual recruitments. There was some conflict between the signals given by the abundance indices and age/length compositions.

The models used were probably over-parameterised considering the limited data available. For example, while sex transition was considered likely for tilefish, the data were not able to estimate the rate that this occurs. SS3 tends to encourage over-parameterisation in models.

Abundance index standardisation was considered part of the data workshop's terms of reference, but perhaps should become part of the stock assessment. It was not clear that the abundance indices and length and age composition measurement errors were estimated in a consistent way. Although the weights given to these sources of information can be adjusted independently, there is often no generally-accepted way to do this without some subjectivity.

The model diagnostics produced as part of the assessments were generally good. Excellent information was provided on how the likelihood changed among components for each run, as well as retrospective analyses, fits and residuals for the length/age compositions and abundance indices. Other useful diagnostics that should be provided routinely are:

1. Residuals for age and length compositions against time.
2. Parameter correlation matrix for the stock assessment model and abundance index standardisation model (or absolute correlations exceeding 0.9 for large numbers of parameters).

The sensitivity runs were extensive and cover most uncertainty. Only one further run was identified as necessary by the review panel. However, it was unclear how sensitivity runs might be used in the management advice. Given that allowable catches will now be adjusted to the risk of overfishing, combining uncertainty into a single coherent form is important.

The main way to measure uncertainty is to carry out Markov Chain Monte Carlo simulations (MCMC). These are a standard way to map and integrate the likelihood (or posterior probability) of the statistical model. MCMC produces frequencies of values of interest (parameters, stock size, stock status indicators etc.) from the underlying density estimate. MCMC frequencies can be combined from two sensitivity runs where they are considered equally probable or combined in proportion to the probability of their being true, assuming they are mutually exclusive. Unfortunately MCMC are time consuming to run and relative probabilities for different models difficult to agree on. It is therefore better to have as few sensitivity runs as possible.

Sensitivity runs, should, where possible be reserved for model structure issues. These would include, but not be limited to, weights applied to different model likelihood components, dividing stocks into different areas, sex specific models with or without sex transition, selectivity and catchability time blocks or change over time (e.g. a random walk process) and so on.

In these assessments, a number of sensitivities were used to profile important parameters such as steepness or natural mortality. These sensitivities could have been converted to input priors, which would then be integrated using a single MCMC. For example, alternative values for natural mortality were provided by the Data Workshop for both stocks and could have been combined either using kernel smoothers or by fitting a parametric probability density function (e.g. normal) to those values. These priors may have to be adjusted to provide the required results (e.g. tested by looking at the likelihood profile on the parameter).

The reference age for the natural mortality could be fixed at a single value following advice from SEDAR 12 (see below), which would avoid unnecessary sensitivities on this parameter. These sensitivity runs only changed the effective natural mortality,

albeit in an obscure way. Therefore, an improved treatment of reference age parameter is set out below.

Natural Mortality Reference Age

Stock Synthesis 3 implements size dependent natural mortality through the model suggested by Lorenzen (2005). As well as specifying the natural mortality, SS3 also requires a reference age. The reference age ensures that the natural mortality is kept relatively constant while growth parameters are estimated. In this formulation, the natural mortality is kept constant over this age regardless of growth parameters, while the length-dependent mortality parameter, natural mortality-at-unit length (M_1), will vary.

There are two alternative ways to deal with this in the software. Firstly, it might be possible to estimate the M_1 parameter directly and provide it to the software. This would require estimating M_1 rather than M from various methods. This approach may still present problems when growth is estimated within the stock assessment, as is the case in most SS3 assessments.

Secondly, a previous review (SEDAR 12) suggested that the M_1 should be set such that the total survival over the main ages selected by the fishery should be the same as that obtained with the constant natural mortality. The estimated natural mortality is therefore an estimate of the average over this period. This is my preferred option.

If SS3 continues to use a single age as reference, defining that reference age might be improved. One of the problems with using a single year as the reference year rather than a range is the model may pivot the natural mortality through this year, producing wider ranges of effective natural mortality over age than expected if growth parameters are poorly estimated. This effect may not be detected as it will occur within the model.

The detailed implementation in SS3 of length-specific natural mortality was not available. I assume that the mortality-at-length was set to equal the required mortality at the reference age. The mortality can be linked to mortality-at-length by (Lorenzen 2005):

$$M = -\frac{M_1}{L_\infty K t}\mathrm{Ln}\left(\frac{L_a}{L_a + L_\infty\left(e^{Kt}-1\right)}\right) \qquad (1)$$

Where M =constant natural mortality, M_1=natural mortality-at-unit-length, L =asymptotic length, K=instantaneous growth rate, L_a=length at the start of the selected period and t=time over the period (age range). This model applies to the von Bertalanffy growth model. The reference age decides the length at the beginning of this period:

$$L_a = L_\infty\left(1 - e^{-Kt_a}\right) \qquad (2)$$

Where t_a=the reference age.

In the case of SS3, the time over the period is one year (t=1, t_a=the reference year). The suggestion from SEDAR 12, to base the natural mortality as the average over the selected ages, is clearly a better way to reference the mortality. I would advise a change to SS3 if possible to an age range (Equation 1) rather than single reference age. This should be a simple alteration to the model.

In the meantime, it is possible to convert an age range to a single reference age by combining equations (1) and (2):

$$-\frac{M_1}{L_\infty Kt}\mathrm{Ln}\left(\frac{L_\infty\left(1-e^{-Kt_a}\right)}{L_\infty\left(1-e^{-Kt_a}\right)+L_\infty\left(e^{Kt}-1\right)}\right)=-\frac{M_1}{L_\infty K}\mathrm{Ln}\left(\frac{L_\infty\left(1-e^{-Kt_b}\right)}{L_\infty\left(1-e^{-Kt_b}\right)+L_\infty\left(e^{K}-1\right)}\right)$$

Where t_a and t_b are the reference ages for the age range (SEDAR 12) and single age (SS3) respectively, and these ages include the integration constant t_0. This can be simplified to:

$$\frac{\left(1-e^{-Kt_a}\right)^{1/t}}{\left(\left(1-e^{-Kt_a}\right)+\left(e^{Kt}-1\right)\right)^{1/t}}=\frac{\left(1-e^{-Kt_b}\right)}{\left(1-e^{-Kt_b}\right)+\left(e^{K}-1\right)}$$

$$t_b=\mathrm{Ln}\left(1-\frac{\left(e^{K}-1\right)\left(1-e^{-Kt_a}\right)^{1/t}}{\left(\left(1-e^{-Kt_a}\right)+\left(e^{Kt}-1\right)\right)^{1/t}-\left(1-e^{-Kt_a}\right)^{1/t}}\right)\Big/K$$

The reference age is a function of the instantaneous growth rate used in the growth model (K), the start age (t_a) and age range (t) to which the natural mortality estimate applies. The start age and age range can often be estimated or is known. Methods based on age for estimating natural mortality would provide this information. These results suggest that the appropriate reference age is not very sensitive to the growth rate.

For yellowedge grouper, the age range used for natural mortality catch curves were 12-41 years (SEDAR 22 DW Final Report). The instantaneous growth rate was between 0.05-0.10 year^{-1}, which would suggest a reference age of 22 or 23 years, with 23 years being more appropriate for the slower growth. The default reference age used was 15 years, which would imply a faster growth or a lower average natural mortality than is the case. It should be noted that the scoping and sensitivity analyses covered these values.

For golden tilefish, the age range used for natural mortality appears to be 4-40 years (it is not stated clearly in the SEDAR 22 DW Final Report). The instantaneous growth rate was estimated to be around 0.10 -0.20 year^{-1}, which would suggest a reference age of 12 to 15 years, with 15 years being more appropriate for the slower growth. The assessment report suggests that the natural mortality was fixed for age 4. An adjusted lower natural mortality and higher reference age might have been more appropriate.

3 Recommend appropriate estimates of stock abundance, biomass, and exploitation.

Additional work is still being conducted and therefore the consensus report is incomplete at this time. The value estimates of abundance, biomass and exploitation will be in the Assessment Workshop report and the Review Workshop Final Report.

In the first instance, six sensitivity runs were selected for each assessment to provide deterministic output indicating a range of results. In each case, of these six, three were selected for full MCMC analysis. The MCMC provides a full analysis of the within-model uncertainty. As indicated above (ToR 3), reducing the number of sensitivity runs in ways other than selection by the review panel would be preferable and should produce more reliable results for risk analysis.

For yellowedge grouper, the main sensitivity runs were the central "base" run, a lower natural mortality and an alternative weighting balancing the influence of the age/length compositions and abundance indices. I suspect that the alternative weighting might represent a better "base" (i.e. most likely), but without further research it is not possible to be certain how to deal with balancing these different sources of information.

For golden tilefish, the main sensitivity runs were the central "base" run, a lower natural mortality and a higher recruitment deviation variance. The model applying an alternative weighting as used for yellowedge, appeared to show problems in model fitting, so was rejected as probably unreliable. The higher recruitment deviations allowed more flexibility for the model in covering alternative stock recruitment uncertainty which was considered desirable at least for the projections.

These sensitivity runs cover the main uncertainty. It is possible to add further sensitivity runs to these and combine their MCMC estimates for risk analysis. This should be done with care, however, because there will be implicit probabilities applied to the possible outcomes depending on which runs are chosen. The three chosen by the Review Panel are probably most appropriate at least until an update assessment is carried out.

4 Evaluate the methods used to estimate population benchmarks and management parameters (e.g., MSY, OFL, Fmsy, Bmsy, MSST, MFMT, or their proxies); recommend appropriate management benchmarks and provide estimated values for management benchmarks, a range of ABC, and declarations of stock status.

Using estimates of MSY in models often depends on parameters which are poorly estimated, such as steepness. MSY is poorly estimated in these cases and therefore I recommend continued use of MSY proxies.

Choice between $SPR_{30\%}$, $SPR_{40\%}$ or some other reference point includes consideration of risk, which the scientists and reviewers are unable to determine. $SPR_{40\%}$ is more precautionary than $SPR_{30\%}$. If the fishing industry wished to obtain a green-label (such as www.msc.org) to improve their market, a more precautionary reference point ($SPR_{40\%}$) would be easier to justify. Any research, including simulation work, to justify the chosen benchmark would be helpful in this case.

The range of ABC values would depend on the approach to risk calculation. This is still subject to a decision from the management council. The ABC should reflect the allowable risk, so an ABC reflecting the 40% (for example) percentile of the MCMC OFL (where the median is 50%) would represent a 40% chance of overfishing. This would seem a good approach if the MCMC values can be combined across sensitivity runs.

Additional assessment work is still being conducted. The benchmarks and management parameters will be put into the Review Workshop Final Report. However, based on the available assessment results, the yellowedge grouper biomass and fishing mortality is very close to the current benchmark ($SPR_{30\%}$), whereas both golden tilefish biomass and fishing mortality is above the current benchmark. It both cases, catches should be adjusted downward, possibly more sharply in the golden tilefish case, to achieve the management objectives.

5 Evaluate the adequacy, appropriateness, and application of the methods used to project future population status; recommend appropriate estimates of future stock condition (e.g., exploitation, abundance, biomass).

Within the limitations imposed by the management structure, projections were adequate and appropriate, and methods were applied correctly. Projections were done in a standard manner within SS3 and the methods have been widely reviewed and found acceptable.

Without a clear harvest control rule, the projections may not provide enough information for choosing an appropriate allowable catch. The ABC control rules require various choices to be made on acceptable risk, and it is possible the projections will have to be re-run before final decisions can be made. Also, alternative harvest control rules which might reduce fishing mortality before a rebuilding plan is required might guard against management and stock assessment errors. I recommend development of this type of rule.

As additional assessment work is still being conducted, the projections will be put into the Review Workshop Final Report.

6 Evaluate the adequacy, appropriateness, and application of methods used to characterize uncertainty in estimated parameters. Provide measures of uncertainty for estimated parameters. Ensure that the implications of uncertainty in technical conclusions are clearly stated.

The methods used to quantify uncertainty in assessments included Markov Chain Monte Carlo (to make a random draw from the likelihood) and sensitivity analyses, which are standard and appropriate methods. MCMC is a particularly useful way to integrate uncertainty, and more use could be made of this facility (See ToR 2 above).

The review panel attempted to identify sensitivities which bracket the uncertainty, such that the key indicators of interest for stock status and the exploitation rate are almost certainly within the range. This uses the review process to reduce a very complex multidimensional problem to a single dimension which is easier to understand.

It was apparent from the MCMC simulation output that a much longer burn-in time was required (i.e. 25% of the total simulation run) compared to that allowed, as the simulations appeared to converge slowly.

The sex-transition parameter for both models was poorly estimated. Poorly estimated parameters usually converge very slowly in MCMC and therefore also their uncertainty may be poorly estimated. The final results were demonstrated to be insensitive to this parameter, so until it can be estimated within the models, it may be better to fix it at some reasonable value.

7 Ensure that stock assessment results are clearly and accurately presented in the Stock Assessment Report and that reported results are consistent with Review Panel recommendations.

As far as possible this term of reference will be met by the Review Panel. The final reports are currently being produced.

8 Evaluate the SEDAR Process as applied to the reviewed assessments and identify any Terms of Reference which were inadequately addressed by the Data or Assessment Workshops.

The SEDAR process is an excellent, if expensive, approach to ensuring that the scientific advice is the best available. The SEDAR process had been applied correctly in the assessment of these stocks. Some tilefish Assessment Workshop terms of reference had not been applied by the time of the review workshop because the assessment team was essentially requesting decisions on how to proceed. This required deciding on some important assumptions to allow the assessment to be completed. These decisions were taken and all terms of reference were completed.

An improvement could be made to the Data Workshop Terms of Reference to help deal with conflicting signals from different information sources. Given that these were relatively data poor fisheries, direct conflict between various sources of information was weak, so simple weighting schemes have provided an adequate solution. However, stronger conflicts may still occur between data sources, and averaging among them may be a poor way to deal with this conflict.

An alternative is for the Data Workshop to provide a semi-quantitative (e.g. a score of 1-5) or quantitative estimate of the relative reliability of each information source. This information could then be used to inform the choice of weights to be used in stock assessments or whether to reject or accept particular sources which are in conflict. This relative reliability given before the assessment and therefore independent of the analysis would be useful. For example, if two abundance indices indicate opposite trends after allowing for selectivity, the base assessment should be chosen assuming one is true. The average between the two trends cannot be correct. DW Terms of reference could include “Provide an index of reliability for each data source, with reference to its interpretation within a stock assessment, based on expert opinion of the DW panel”.

Some improvements may be made to the Review Terms of Reference (ToR). Given SS3 will be used, many of the ToRs referring to methods may become redundant. ToRs 4 and 5 require evaluating methods which are standard approaches, have already been reviewed extensively and been found acceptable.

The ToRs requiring appropriate values and estimates derived from the stock assessment also appear redundant for the independent CIE reports. Final versions of these values were not available in time for this report and these values summarising the results of the stock assessment will be reported in the consensus review report anyway. Statements indicating the status of the stock and to the effect that the stock assessment provides the best available scientific advice seem reasonable and necessary.

The review ToRs should focus on requiring the best possible stock assessment, taking into account limitations on data, methods and resources available, but recognise limitations need to be commensurate with the size and level of exploitation of the fishery. Risks of overfishing can be reduced by improving information or by reducing the exploitation level. Where changes are required by reviewers, these must recommend specific practical changes that can be carried out within a reasonable time frame. In general, it should be sufficient for the SEDAR process to demonstrate an

ongoing feedback system that will produce improving assessments, rather than a perfect assessment on a single iteration. It should only be possible to reject an assessment when a clear and correctable error can be identified and a solution provided.

9 Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted. Clearly denote research and monitoring needs that could improve the reliability of future assessments. Recommend an appropriate interval for the next assessment, and whether a benchmark or update assessment is warranted.

All research suggested by the Data and Assessment Workshops seem appropriate and should lead to improved stock assessments. In addition, I have made some suggestions below which might provide inexpensive solutions to some issues which have arisen during the review.

Criteria for choosing what research needs to be done needs to be decided on the basis of cost, likely success of the research and impact of the research on the outcome of the assessment. The main areas of research recommendations broadly divide into more data collection and improvements in methodology. While more important in the long term, improved data collection may prove too difficult or expensive to achieve in the near term. In contrast, improvements in methods should be relatively inexpensive and have an impact in the short-term.

I believe that initial research should focus on balancing weights between the different sources of information and combining information into as convenient a form as possible for decision-making. These ideas, also outlined in other parts of this report, include:

- Combine sensitivity runs using more priors over key parameters such as natural mortality (M).
- Improve the standardisation GLM for abundance indices, with the objective of reducing abundance index error and ensuring remaining errors are scaled appropriately in comparison with the age/length compositions.

In addition, there may be a significant improvement in the assessments if the standardisation were included within the stock assessment model, so that a genuine weighting among different data sets can be applied.

One of the strengths of SS3 is that the model is fitted to raw data rather than processed values. This allows the model to use likelihoods which have a sound theoretical basis. In reality, this is only true for the age and length compositions. The abundance indices are usually heavily processed through a standardisation model, but could be provided as raw catch and effort with standardising covariates which might be used to allow for changes in catchability.

Having a GLM as part of the assessment model could get complicated. As a trial, a very simple biomass dynamics model incorporating a GLM with simulated data might be used to see how standardisation could affect the assessment. If incorporating standardisation in the assessment appeared useful, a more complex model incorporating a simple GLM into an age structured assessment using a trial

configuration in ADMB, following as far as possible the SS3 model, could be ventured. This work could feed into improving SS3 providing alternative options.

There are two issues to consider in deciding on when the next assessment should take place. Firstly, if the stock is close to being overfished, frequent update assessments may be required. On the other hand, if new data become available, management changes (harvest control rule, the targets or limits) or new methods are developed, a full benchmark assessment should be undertaken. In any case, the maximum time before an update should be five years and maximum time before a benchmark should be 10 years to ensure best possible advice is available. Given the state of the stocks, annual update assessments are justifiable at least until overfishing is brought to the target level. Based on these criteria, both golden tilefish and yellowedge grouper should have an update after two years, because they are at least fully exploited. If there are significant improvements in the available methods, update assessments could be upgraded to benchmark assessments.

10 Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Develop a list of tasks to be completed following the workshop.

Most of the Peer Review Summary report has been completed, although it awaits some additional work to be completed by the assessment teams. This work mainly consists of completing the MCMC analysis for each of the three sensitivity runs identified for each assessment (see ToR 3).

Conclusions and Recommendations

The data and assessments reported by the review panel form a good basis for scientific advice. Uncertainties associated with the assessment are dealt with appropriately and taken into account. The results represent the best scientific advice available for these stocks.

All data were appropriate, and suitable for use in stock assessment. The main weakness identified was the treatment of the abundance indices. Suggested improvements were alternative, more parsimonious configurations for the standardising GLM and including standardisation within the stock assessment model.

The main assessment software, Stock Synthesis 3 (SS3), is appropriate for the type of data available and should enable the scientists to set up assessments and obtain credible and consistent results.

Additional work was still being conducted before the completion of this report. The consensus Review Workshop Final Report will have a summary of final assessment results. Based on the available assessment results, the yellowedge grouper biomass and fishing mortality is very close to the current benchmark (SPR30%), whereas golden tilefish biomass and fishing mortality is above the current benchmark. It both cases, catches should be adjusted downward, possibly more sharply in the golden tilefish case, to achieve the management objectives.

The SS3 parameter, reference age of the natural mortality-at-length, should be fixed based on the relevant ages provided for the natural mortality estimate. The way this might be done is described under ToR 2.

The large number of sensitivity runs to scope the uncertainty in the assessment presents a problem for combining this uncertainty into a single form for advice. I suggest reducing sensitivity runs as much as possible, though developing priors for key parameters, such as natural mortality, rather than a series of fixed values over a range.

References

Lorenzen, K. (2005) Population dynamics and potential of fisheries stock enhancement: practical theory for assessment and policy analysis. Philosophical Transactions of the Royal Society of London. Fisheries Theme Issue 2004

McCullagh P, Nelder JA (1989) Generalized linear models. Second Edition. Chapman and Hall, New York.

Appendix I: *SEDAR 22* Workshop Document List

Gulf of Mexico Yellowedge Grouper and Tilefish
Workshop Document List

Document #	Title	Authors	Working Group
Documents Prepared for the Data Workshop			
SEDAR22-DW-01	Golden tilefish (*Lopholatilus chamaeleonticeps*) age, growth, and reproduction from the northeastern Gulf of Mexico: 1985,1997-2009	Linda Lombardi, Gary Fitzhugh, Hope Lyon	Life History
SEDAR22-DW-02	Commercial longline vessel standardized catch rates of yellowedge grouper in the Gulf of Mexico	Neil Baertlein and Kevin McCarthy	Indices
SEDAR22-DW-03	Golden tilefish and blueline tilefish standardized catch rates from commercial longline vessels in the Gulf of Mexico	Kevin McCarthy	Indices
SEDAR22-DW-04	Discards of yellowedge grouper, golden tilefish, and blueline tilefish from commercial fishing vessels in the Gulf of Mexico	Kevin McCarthy	Catch Statistics
SEDAR22-DW-05	Explorations of habitat associations of yellowedge grouper and golden tilefish	John F Walter, Melissa Cook, Brian Linton, Linda Lombardi, and John A. Quinlan	Life History
SEDAR22-DW-06	Abundance Indices of subadult Yellowedge Grouper, *Epinephelus flavolimbatus*, Collected in Summer and Fall Groundfish Surveys in the northern Gulf of Mexico	Adam G. Pollack and G. Walter Ingram, Jr.	Indices
SEDAR22-DW-07	Abundance Indices of Yellowedge Grouper and Golden Tilefish Collected in NMFS Bottom Longline Surveys in the northern Gulf of Mexico	G. Walter Ingram, Jr. and Adam G. Pollack	Indices
SEDAR22-DW-08	Yellowedge grouper (Epinephelus flavolimbatus) age, growth and reproduction from the northern Gulf of Mexico	Melissa Cook and Michael Hendon	Life History
SEDAR22-DW-09	Observed Length frequency distributions and otolith sampling	Ching-Ping Chih	Life History/ Catch

	issues for yellowedge groupers caught in the Gulf of Mexico from 1984 to 2009.		Statistics
SEDAR22-DW-10	Observed Length frequency distributions and otolith sampling issues for tile fish caught in the Gulf of Mexico from 1984 to 2009	Ching-Ping Chih	Life History/ Catch Statistics
SEDAR22-DW-11	Length frequency distributions for blue line tile fish caught in the Gulf of Mexico from 1984 to 2009	Ching-Ping Chih	Life History/ Catch Statistics
SEDAR22-DW-12	Estimation of species misidentification in the commercial landing data of tile fish in the Gulf of Mexico from 1984 to 2009	Ching-Ping Chih	Catch Statistics
SEDAR22-DW-13	Estimation of species misidentification in the commercial landing data of yellowedge groupers in the Gulf of Mexico from 1984 to 2009	Ching-Ping Chih	Catch Statistics
SEDAR22-DW-14	Evidence of hermaphroditism in Golden Tilefish (*Lopholatilus chamaeleonticeps*) in the Gulf of Mexico	Hope Lyon	Life History
SEDAR22-DW-15	Recreational Survey Data for Yellowedge Grouper, Tilefish (golden), and Blueline Tilefish in the Gulf of Mexico	Vivian M. Matter	Catch Statistics
SEDAR22-DW-16	Estimated Recreational Catch in Weight: Method for Filling in Missing Weight Estimates from the Recreational Surveys	Vivian M. Matter	Catch Statistics
SEDAR22-DW-17	Commercial Landings of Yellowedge Grouper, Golden Tilefish, and Blueline Tilefish from the Gulf of Mexico region	Refik Orhun	Catch Statistics
Documents Prepared for the Assessment Process			
SEDAR22-AP-01	United States Commercial Longline Vessel Standardized Catch Rates of Golden and Blueline Tilefish in the Gulf of Mexico, 1992-2009: Revised	Kevin McCarthy	
SEDAR22-AP-02	United States Commercial Longline Vessel Standardized Catch Rates of Yellowedge Grouper (*Epinephelus flavolimbatus*) for Three Regions in the Gulf of Mexico, 1991-2009	Neil Baertlein and Kevin McCarthy	

SEDAR22-AP-03	Pre-review draft of the tilefish assessment report (23 Nov 2010)	
SEDAR22-AP-04	Pre-review draft of the yellowedge grouper assessment report (23 Nov 2010)	
Documents Prepared for the Review Workshop		
SEDAR22-RW-01		
Final Stock Assessment Reports		
SEDAR19-SAR1	Yellowedge Grouper	
SEDAR19-SAR2	Golden Tilefish	
Reference Documents		
SEDAR22-RD01	Lead-radium dating of golden tilefish (*Lopholatilus chamaeleonticeps*)	Allen Andrew
SEDAR22-RD02	Status of the yellowedge grouper fishery in the Gulf of Mexico	Shannon L. Cass-Calay and Melissa Bahnick
SEDAR22-RD03	Yellowedge grouper (*Epinephelus flavolimbatus*) and golden tilefish (*Lopholatilus chamaeleonticeps*) distributions, habitat preferences and available biological samples	Melissa Cook and Linda Lombardi-Carlson
SEDAR22-RD04	Validation of yellowedge grouper, *Epinephelus flavolimbatus*, age using nuclear bomb-produced radiocarbon	Melissa Cook & Gary R. Fitzhugh & James S. Franks
SEDAR22-RD05	Population dynamics structure, and per –recruit analyses of yellowedge grouper, *Epinephelus flavolimbatus* from the northern Gulf of Mexico	Melissa Cook
SEDAR22-RD06	Reproduction of yellowedge grouper *Epinephelus flavolimbatus*, from the eastern Gulf of Mexico	Bullock, L. H., M. F. Godcharles and R. E. Crabtree
SEDAR22-RD07	Burrow utilization by yellowedge grouper, *Epinephelus flavolimbatus*, in the northwestern Gulf of Mexico	Jones, R. S., E. J. Gutherz, W. R. Nelson and G. C. Matlock
SEDAR22-RD08	Age and growth of the yellowedge grouper, *Epinephelus flavolimbatus*, and the yellowmouth grouper, *Mycteroperca interstitialis*, off Trinidad and Tobago	Manickchand-Heileman, S. C. and D. A. T. Phillip
SEDAR22-RD09	A descriptive survey of the bottom longline fishery in the Gulf of Mexico	Prytherch, H. F.
SEDAR22-RD10	Comparison of Two Techniques for	Matlock, Gary C., Walter R.

	Estimating Tilefish, Yellowedge Grouper, and Other Deepwater Fish Populations	Nelson, Robert S. Jones, Albert W. Green, Terry J. Cody, Elmer Gutherz, and Jeff Doerzbacher
SEDAR22-RD11	Deep-water sinkholes and biotherms of South Florida and the Pourtales Terrace – Habitat and Fauna	John K. Reed, Shirley A. Pomponi, Doug Weaver, Charles K. Paull, and Amy E. Wright
SEDAR22-RD12	Tilefishes of the genus *Caulolatilus* construct burrows in the sea floor	K.W. Able, D.C. Twichell, C.B. Grimes, and R.S. Jones
SEDAR22-RD13	Spawning Locations for Atlantic Reef Fishes off the Southeastern U.S.	GEORGE R. SEDBERRY, O. PASHUK, D.M. WYANSKI, J.A. STEPHEN, and P. WEINBACH
SEDAR22-RD14	Trends in tilefish distribution and relative abundance off South Carolina and Georgia	Charles A. Barnes and Bruce W. Stender
SEDAR22-RD15	Age, growth, and reproductive biology of blueline tilefish along the Southeastern coast of the United States, 1982-1999	Patrick J. Harris, David M. Wyanski, and Paulette T. Powers Mikell
SEDAR22-RD16	Temporal and spatial variation in habitat characteristics of tilefish (Lopholatilus chamaeleonticeps) off the east coast of Florida	Kenneth W. Able, Churchill B. Grimes, Robert S. Jones and David C. Twichell
SEDAR22-RD17	The Complex Life History of Tilefish *Lopholatilus chamaeleonticeps* and Vulnerability to Exploitation	Churchill B. Grimes and Stephen C. Turner
SEDAR22-RD18	The fishery for tilefish, *Lopholatilus chamaeleonticeps*, off South Carolina and Georgia	Bob Low, Glenn Ulrich, and Frank Blum
SEDAR22-RD19	Tilefish off South Carolina and Georgia	R.A. Low, Jr., G.F. Ulrich, and F. Blum
SEDAR22-RD20	Spawner-recruit relationships of demersal marine fishes: Prior distribution of steepness for possible use in SEDAR stock assessments	SEDAR 24−AW−06 - Sustainable Fisheries Branch

Appendix II: Statement of Work for Dr. Paul Medley

External Independent Peer Review by the Center for Independent Experts

SEDAR 22 Gulf of Mexico Yellowedge Grouper and Tilefish 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 22 will be a compilation of data, a benchmark assessment of the stock, and an assessment review conducted for Gulf of Mexico Yellowedge Grouper and Tilefish. The review workshop provides an independent peer review of SEDAR stock assessments. The review panel is ultimately responsible for ensuring that the best possible assessment is provided through the SEDAR process. The stocks assessed through SEDAR 22 are within the jurisdiction of the Gulf of Mexico Fishery Management Council and the states of Texas, Louisiana, Mississippi, Alabama, and Florida. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the review meeting is 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 expertise, working knowledge, and recent experience in stock assessment, statistics, fisheries science, and marine biology sufficient to complete the primary task of reviewing the technical details of the methods used for the assessment. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Tampa, Florida during 14-17 February 2011.

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.

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

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

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

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

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

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting in Tampa, Florida during 14-17 February 2011.
3) During 14-17 February 2011 in Tampa, Florida as specified herein, conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than 3 March 2011, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and CIE Regional Coordinator, via email to David Sampson david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

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

11 January 2011	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
1 February 2011	NMFS Project Contact sends the CIE Reviewers the pre-review documents
14-17 February 2011	Each reviewer participates and conducts an independent peer review during the panel review meeting
3 March 2011	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
17 March 2011	CIE submits CIE independent peer review reports to the COTR
24 March 2011	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: Requests to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on substitutions. 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) each CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) each 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:

William Michaels, Contracting Officer's Technical Representative (COTR)
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

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

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

Key Personnel:

NMFS Project Contact:

Julie A Neer, SEDAR Coordinator

4055 Faber Place Drive, Suite 201, North Charleston, SC 29405

Julie.neer@safmc.net Phone: 843-571-4366

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: Tentative Terms of Reference for the Peer Review

SEDAR 22 Gulf of Mexico Yellowedge Grouper and Tilefish Review Workshop

Yellowedge Grouper:

1. Evaluate the adequacy, appropriateness, and application of data used in the assessment.
2. Evaluate the adequacy, appropriateness, and application of methods used to assess the stock.
3. Recommend appropriate estimates of stock abundance, biomass, and exploitation.
4. Evaluate the methods used to estimate population benchmarks and management parameters (*e.g., MSY, OFL, Fmsy, Bmsy, MSST, MFMT, or their proxies*); recommend appropriate management benchmarks and provide estimated values for management benchmarks, a range of ABC, and declarations of stock status.
5. Evaluate the adequacy, appropriateness, and application of the methods used to project future population status; recommend appropriate estimates of future stock condition (e.g., exploitation, abundance, biomass).
6. Evaluate the adequacy, appropriateness, and application of methods used to characterize uncertainty in estimated parameters. Provide measures of uncertainty for estimated parameters. Ensure that the implications of uncertainty in technical conclusions are clearly stated.
7. Ensure that stock assessment results are clearly and accurately presented in the Stock Assessment Report and that reported results are consistent with Review Panel recommendations.
8. Evaluate the SEDAR Process as applied to the reviewed assessments and identify any Terms of Reference which were inadequately addressed by the Data or Assessment Workshops.
9. Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted. Clearly denote research and monitoring needs that could improve the reliability of future assessments. Recommend an appropriate interval for the next assessment, and whether a benchmark or update assessment is warranted.
10. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Develop a list of tasks to be completed following the workshop.

The review panel may request additional sensitivity analyses, evaluation of alternative assumptions, and correction of errors identified in the assessments provided by the assessment workshop panel; the review panel may not request a new assessment. Additional details regarding the latitude given the review panel to deviate from assessments provided by the

assessment workshop panel are provided in the *SEDAR Guidelines* and the *SEDAR Review Panel Overview and Instructions*.

** The panel shall ensure that corrected estimates are provided by addenda to the assessment report in the event corrections are made in the assessment, alternative model configurations are recommended, or additional analyses are prepared as a result of review panel findings regarding the TORs above.**

Tilefish:

1. Evaluate the adequacy, appropriateness, and application of data used in the assessment.
2. Evaluate the adequacy, appropriateness, and application of methods used to assess the stock.
3. Recommend appropriate estimates of stock abundance, biomass, and exploitation.
4. Evaluate the methods used to estimate population benchmarks and management parameters (*e.g., MSY, OFL, Fmsy, Bmsy, MSST, MFMT, or their proxies*); recommend appropriate management benchmarks and provide estimated values for management benchmarks, a range of ABC, and declarations of stock status.
5. Evaluate the adequacy, appropriateness, and application of the methods used to project future population status; recommend appropriate estimates of future stock condition (e.g., exploitation, abundance, biomass).
6. Evaluate the adequacy, appropriateness, and application of methods used to characterize uncertainty in estimated parameters. Provide measures of uncertainty for estimated parameters. Ensure that the implications of uncertainty in technical conclusions are clearly stated.
7. Ensure that stock assessment results are clearly and accurately presented in the Stock Assessment Report and that reported results are consistent with Review Panel recommendations.
8. Evaluate the SEDAR Process as applied to the reviewed assessments and identify any Terms of Reference which were inadequately addressed by the Data or Assessment Workshops.
9. Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted. Clearly denote research and monitoring needs that could improve the reliability of future assessments. Recommend an appropriate interval for the next assessment, and whether a benchmark or update assessment is warranted.
10. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Develop a list of tasks to be completed following the workshop.

The review panel may request additional sensitivity analyses, evaluation of alternative assumptions, and correction of errors identified in the assessments provided by the assessment workshop panel; the review panel may not request a new assessment. Additional details regarding the latitude given the review panel to deviate from assessments provided by the

assessment workshop panel are provided in the *SEDAR Guidelines* and the *SEDAR Review Panel Overview and Instructions*.

** The panel shall ensure that corrected estimates are provided by addenda to the assessment report in the event corrections are made in the assessment, alternative model configurations are recommended, or additional analyses are prepared as a result of review panel findings regarding the TORs above.**

Annex 3: Tentative Agenda

SEDAR 22 Gulf of Mexico Yellowedge Grouper and Tilefish Review Workshop

Tampa, Florida
14-17 February 2011

Monday

1:00 p.m.	**Convene**	
1:00 – 1:30	**Introductions and Opening Remarks** *- Agenda Review, TOR, Task Assignments*	**Coordinator**
1:30 – 3:30	**Assessment Presentation**	**TBD**
3:30 – 4:00	**Break**	
4:00 – 6:00	**Continue Presentation/Discussion**	**Chair**

Tuesday

8:30 a.m. – 11:30 a.m.	**Assessment Presentation**	**Chair**
11:30 a.m. – 1:30 p.m.	**Lunch Break**	
1:30 p.m. – 3:30 p.m.	**Panel Discussion** **-** *Assessment Data & Methods* *- Identify additional analyses, sensitivities, corrections*	**TBD**
3:30 p.m. – 4:00 p.m.	**Break**	
4:00 p.m. – 6:00 p.m.	**Panel Discussion** *- Continue deliberations* *- Review additional analyses*	**Chair**

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

Wednesday

8:30 a.m. – 11:30 a.m.	**Panel Discussion** *- Review additional analyses, sensitivities* *- recommendations and comments*	**Chair**
11:30 a.m. – 1:30 p.m.	**Lunch Break**	
1:30 p.m. – 3:30 p.m.	**Panel Discussion**	**TBD**
3:30 p.m. – 4:00 p.m.	**Break**	
4:00 p.m. – 6:00 p.m.	**Panel Discussion**	**Chair**

Wednesday Goals: Final sensitivities identified, Preferred models selected, Projection approaches approved, Report drafts begun

Thursday

8:30 a.m. – 11:30 a.m.	**Panel Discussion** *- Final sensitivities reviewed.* *- Projections reviewed.*	**Chair**
11:30 a.m. – 1:30 p.m.	**Lunch Break**	
1:30 p.m. – 3:30 p.m.	**Panel Discussion or Work Session**	**Chair**
3:30 p.m. - 4:00 p.m.	**Break**	
4:00 p.m. - 6:00 p.m.	**Panel Work Session** *- Review Reports*	**Chair**

Thursday Goals: Complete assessment work and discussions. Final results available. Draft Reports reviewed.