# Individual CIE reviewer's report on the 2011 South East Data, Assessment and Review (SEDAR) 25, review workshop on South Atlantic black sea bass and golden tilefish stocks, 11th-13th October 2011, Charleston, South Carolina

Michael T. Smith[1],
Centre for Fisheries, Environment & Aquaculture Science (Cefas),
Pakefield Road,
Lowestoft,
Suffolk, NR33 0HT
United Kingdom.
mike.smith@cefas.co.uk

[1]Representing the Center for Independent Experts

# Contents

## 1. Executive Summary

This document is the individual CIE Reviewer report of the 2011 South East Data, Assessment and Review (SEDAR 25), Review Workshop on South Atlantic black sea bass and golden tilefish stocks.

The review took place on 11-13th October 2010 in Charleston, South Carolina. Material presented was derived from separate Data and Assessment Workshops. The Review Panel comprised the Chair, two South Atlantic Science & Statistical Committee representatives and three reviewers appointed by the CIE. The review was also attended by the SEDAR coordinator, NMFS and fishery committee representatives as well as members of the public and fishing industry. The Review Panel Chair ran the meeting efficiently and cordially and the process was effectively supported by the SEDAR Coordinator. Contributions by members of the industry and SAFMC members were very useful in providing fishery and management context for the external reviewers.

For both stocks, the primary assessment model was the Beaufort Assessment Model (BAM), a statistical age-structured model, capable of synthesising time-series of data including: total landings and discards, length and age compositions and abundance indices. Monte Carlo bootstrap (MCB) and sensitivity runs were used to evaluate uncertainty in results and a surplus production model, ASPIC was used to provide an alternative assessment assuming different population dynamics. This framework provided a very effective and flexible tool for assessment, projection and consideration of uncertainty.

Black sea bass

The black sea bass assessment indicated that F was above $F_{MSY}$ and although SSB has increased above MSST it has not yet reached $SSB_{MSY}$, the rebuilding target. This result was generally consistent over a range of uncertainty in data, model configurations and population dynamics assumptions. The stock recruitment relationship caused some concern with temporal structuring apparent in the residuals. Whether these reflected a real effect or an artefact was not clear, but this merits further investigation, initially via analysis of SR fits in the MCB analysis, which if available in the assessment outputs could indicate whether or not the effect is data-driven. The major source of uncertainty in this assessment was considered to relate to monitoring of the recreational fishery which forms the major component of the landings. Improvements to monitoring programmes for the recreational sector would therefore be beneficial, although maintaining consistency with historical time-series is also crucial.

Golden tilefish

The golden tilefish assessment indicated that the stock is currently neither overfished nor suffering overfishing. This result was generally consistent across a range of uncertainty in the data, model configurations and population dynamics assumptions. The stock–recruitment relationship was one of the major sources of uncertainty because the steepness parameter could not be estimated and a single large residual also gave some cause for concern. The lack of a reliable fishery-independent index was also considered a significant source of uncertainty. Knowledge and data relating to the golden tilefish stock are lacking with poor quantification of life-history parameters, relatively low levels of biological sampling and difficulties in age determination. There is substantial scope to carry out further research on this species if this can be politically or economically justified.

## 2. Background

The South East Data, Assessment, and Review (SEDAR) is a cooperative process for conducting and peer-reviewing stock assessments for stocks off the US southeastern seaboard, Gulf of Mexico and the Caribbean. It is supported by the South Atlantic, Gulf of Mexico and Caribbean Fishery Management Councils, NOAA Fisheries, SEFC, SERO and the Atlantic and Gulf States Marine Fisheries Commissions. The SEDAR process consists of separate data, assessment and review workshops.

The previous assessment of South Atlantic black sea bass was conducted in 2002/3 (SEDAR 02) and updated in 2005/6, using a similar form of model to the one implemented at SEDAR 25. Golden tilefish were assessed previously in SEDAR 04 (2006) as part of the south Atlantic snapper/grouper complex. Data for SEDAR 25 were compiled during the Data Workshop (DW), and assessment and projection models developed and applied at the Assessment Workshop (AW). The data and assessment workshop reports along with nine working documents were provided to the October 2011 Review Workshop.

## 3. Individual Reviewer's Role in the Review Activities

Review documentation, including data and assessment workshop reports and review working documents (see Bibliography, Appendix 1), were provided to reviewers in electronic format adequately in advance of the review meeting, although a revised version of the golden tilefish Assessment Report and supporting Working Paper (RW06) was uploaded one week prior to the meeting, rather than two weeks as promised in the ToR. Prior to the review meeting, these documents were read and pre-reviewed against the terms of reference (ToR), as specified by the CIE (see Statement of work, Appendix 2). Many additional background documents were also provided.

The SEDAR 25 Review Workshop (RW) took place at the Crowne Plaza Conference Center in Charleston, South Carolina, from 09:00 Tuesday 11th October to 13:00 Thursday 13th October 2011. The provisional agenda for the meeting is given in Annex 3 of Appendix 2.

The Review Panel comprised the Chair, two South Atlantic Science and Statistical Committee representatives and three reviewers appointed by the CIE (see Panel membership, Appendix 3). The RW was also attended by the SEDAR coordinator, NMFS and fishery committee representative, as well as members of the public and industry.

Prior to the meeting, reviewers were asked to volunteer to draft text for particular sections of the report for each species. In my case these were sections 5 and 6, on projections and characterising uncertainty, respectively. Although reviewers would focus slightly more on their own volunteered aspects, they would nonetheless still contribute to the review on all aspects of data compilation and assessment.

Background data and assessment results were presented clearly and concisely by the lead assessment scientists from the AW. Clarification was provided by the assessment presenters and also by other attendees at the meeting. Contributions to discussions by members of industry were particularly appreciated by the external reviewers with regards to providing useful background on the fisheries, and inputs from SAFMC members contributed greatly to setting the management context. The Review Panel requested a number of additional projections during the meeting, which were completed quickly and effectively by the assessment teams and helped to clarify the issues concerned.

The Review Panel Chair ran the meeting efficiently and cordially and the process was effectively supported by the SEDAR Coordinator. The Review Panel provided a Summary Report that reflects the consensus of views reached. This report summarises my own views, which also underpin this individual CIE Reviewer report. I therefore accept responsibility for any errors caused by my misinterpretation of the data or analyses.

# 4. Summary of Findings for each ToR

The review meeting terms of reference are specified in Appendix II, Annex 2, and are identical for both species. Here, they are considered separately by species.

## Black Sea Bass

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

The data workshop report lists 24 working papers, many of which related to black sea bass and considered catch-rate data from the commercial and recreational fisheries as well as fishery-independent indices and/or their standardization (e.g. DW 02, 03, 08, 12, 13, 14, 16, 18, 23 and 24). Others related to length frequency data, discards and other issues in the fishery. The DW report contains summaries for three of these relating to discard mortalities and batch fecundity, noting that DW10 was deemed pertinent in providing a summary of literature on discard rates for both black sea bass and tilefish. The range of documents submitted provided a good basis for consideration of the data issues for the stock assessment.

***Spatial factors***

Working documents submitted by industry to the review meeting raised their concerns regarding stock and fishery structure for black sea bass. One DW reported on a genetic study (McCartney and Burton, SEDAR25-RD42) undertaken to clarify stock definition, which indicated three major regional stocks with some mixing between Mid-Atlantic and South Atlantic stocks around the Cape Hatteras, North Carolina and Virginia dividing line. Preliminary results from another study, utilising otolith microchemistry, suggested no mixing across the Cape Hatteras boundary. In my opinion, the current stock definitions appear appropriate, given existing knowledge.

The DW also investigated demographic patterns between the Carolinas and Florida, through analysis of mean length-at-age of black sea bass in the commercial and recreational fisheries, concluding that there were no discernible differences. Again, this suggests that the spatial scale used for assessment is appropriate. However, some additional consideration of the spatial structure of exploitation, showing areas of high and low (e.g. closed areas) exploitation in relation to the stock distribution and suitable habitat, would be useful to highlight areas with the potential for localised depletion.

Information on movements was available from five tagging studies, all of which tended to suggest high site fidelity on offshore reefs and generally very limited movements. These provided no direct evidence in conflict with the scale used for assessment. However, fishers report migrations of large black sea bass, and juveniles tend to be found inshore. The DW indicated that research into migration and dispersal patterns and their timing is required, and this recommendation is supported because it would help to clarify spatial issues relating to local effects within the overall stock area.

***Life history parameters***

The updated age data for this assessment included >60 000 newly processed samples from both fishery-independent and -dependent samples, with enhanced data on translucent edge presence and month of capture, permitting more precise age determination. Growth parameters were revised, adjusting fishery-dependent data for temporal changes in minimum size regulations.

Point estimates of natural mortality were calculated using a wide range of alternative formulations and requiring some additional life history parameters ($L_\infty$, k, age at maturity, maximum age). The $\text{Hoenig}_{\text{fish}}$ (1983) estimate was proposed and was relatively central to most of the other parameter estimates. This represented an increase in the rate of natural mortality compared with the previous assessment, but was supported by the data analysis carried out. Natural mortality was structured by age using the method of Lorenzen (1996, 2005), which inversely relates M to mean weight-at-age with a power function and has been widely applied in the region. The Lorenzen method was preferred over the newly published approach of Gislason *et al.* (2010), because the DW felt that the latter required further investigation before its use in SEDAR. The analyses carried out by the DW on natural mortality were well thought out and appropriate given the data available, and represented an improvement over the previous constant natural mortality-at-age assumption. Recommendations for sensitivity runs exploring alternative low and high levels of natural mortality were also appropriate.

Black sea bass are protogenous hermaphrodites, changing sex from female to male. The DW investigated data for >39 000 individual fish for which sex, maturity and age data were available, and with significant new data was able to revise maturity estimates from those of SEDAR 2. Logistic regression was used to estimate ogives for both maturity and sexual transition. A single maturity ogive for the whole time-period, rather than changing over three periods (as in SEDAR 2) was recommended, because the $L_{50}$% maturity showed little change through time. Maturity at age 0 was set to zero. A single sex transition ogive for the whole time-period was also considered appropriate.

Batch fecundity against size and age was investigated in two studies producing similar results, and suggesting that size exhibited the stronger relationship. Although BAM is primarily age-structured it can accommodate length-structured data, and the DW recommended using the relationship with fish weight in the assessment. A constant point estimate of 31 batches per female per year was used to estimate annual egg production.

Revisions to the reproductive cycle were generally well supported by the data and appropriate for the assessment. However, it is possible or indeed likely that sexual transition will be variable through time, possibly in response to densities or proportions of females or males in the population, or in response to environmental variables such as temperature. Further work to explore whether these effects are present and the data can support the estimation of parameters to quantify them would be relevant. This might also include a brief meta-analysis of other species to set the context for such studies. It is also possible that the number of batches spawned per year could be related to size/age or environmental conditions, and further studies to evaluate this would be valuable.

***Discard mortality***

A review of literature noted widely varying rates of discard mortality, but a recent tagging study with high rates of return suggested low discard mortality. Studies suggested that discard mortality increased with depth, particularly at depths greater than those where most of the fishery is prosecuted. Discard mortality for traps was lower than for line-caught black sea bass, and larger-meshed escape panels in traps also resulted in less discard mortality. No discard data were available for trawls. Point estimates of 7%, 5% and 1% for hook & line and escape panels with 1.5” mesh and 2” mesh, respectively, were commended, along with an upper bound equivalent to the SEDAR 2 estimate of 0.15. Estimates for discard mortality were based on recent research and structured to be consistent when uncertainty was applied. Comparison with the previous assessment was provided through a continuity sensitivity run. Discard mortality is likely to be highly variable, dependent on environmental conditions (e.g. depth, temperature), gear and fish condition, as well as on treatment of the fish. The latter may vary widely and potentially more so in recreational fisheries, where the catching sector consists of many people. Discard mortality is also difficult to estimate well, often being confounded by the tagging or containment methods used in studies to evaluate it. In this fishery it was considered low, because most of the fishery is in relatively shallow water, and I support this belief. However, given that some studies indicated much higher discard mortality in deeper water, further investigation into discard mortality would be useful.

***Catch***

General recreational landings made up most of the catch, followed by commercial pots, commercial vertical line and trawl fisheries, and headboats. Landings and discard estimates are available from a range of programmes that vary in duration and extent and quality of information according to the collecting authorities and fleets concerned. The DW report lists decisions made to accommodate landings where stock area or species was not fully specified and to compile time-series of landings and discard information. Where time-series did not cover the full assessment duration, data were modelled but not fitted in BAM. Changes through time in minimum landings size were modelled in the assessment by assigning different time-periods for the selectivity curves estimated. Newly keyed historical data (1975–1977) for the headboat catch were accepted for use, and partial spatial coverage by the headboat survey during this period was adjusted for using 3-year average ratios of landings. Discarded fish are reported by anglers in the recreational sampling programme, but not verified, and the programme also does not record size or weight of discards. Self-reported discard data were available for the headboat fleet since 2004, but data from the recreational observer programme suggested that they were underestimated. A proxy for headboat discards was estimated from MRFSS charterboat data from 1986 to 2010. Compilation of data from different data sources that may have changed sampling protocols through the time-series poses many problems. Discards tend to be generally less visible than landings, and hence more difficult to estimate. Catch data for recreational fisheries are frequently based on intercept surveys supplemented by household surveys to determine total effort and raising factors. The scope for bias and/or errors in such surveys is large. Decisions outlined in the DW and AW reports were made on a rational basis in order to make use of as many of the data as possible while maintaining the integrity of time-series, and this seems largely to have been achieved.

***Age and length distributions***

Biological sampling data were obtained from the NMFS trip interview programme (TIP) undergoing automated quality control during their selection from the database. The DW noted that length sampling was inadequate for gears other than handline and pots and that sample sizes needed to be considered carefully. The assessment model is primarily age-structured, with length-at-age modelled secondarily and permitting the use of length data in the fitting process. Where samples contained both length and age data of adequate quality, only the age data were used, to prevent double usage of the same sample. The use of both age and length data increases the amount of potentially useful information available for assessment, and is recommended. Sensible quality assurance criteria were applied to the data during their extraction and inclusion for fitting.

***Abundance indices***

Many of the working papers submitted to the DW related to abundance index construction, and the DW noted that these were helpful in determining which should be recommended for use in the assessment. Recommended indices included two fishery-independent indices (MARMAP chevron traps, 1990–2010, and MARMAP blackfish and Florida antillean trap combined, 1981–1987), both of which were generated by catches of 90-minute soaks at randomly selected stations and standardised using delta-GLM models. Three fishery-dependent indices were also selected (Headboat, 1979–2010, Commercial line, 1993–2010, and Headboat discards, 2005–2010). The headboat index was standardised using a lognormal GLM, with variance estimates produced by bootstrap, but CVs were considered unrealistically low, so for assessment were scaled to 0.3 prior to 1984 and 0.15 thereafter. Although a fishery-dependent index, the headboat index had good spatial and temporal coverage and selection of trips was for those primarily targetting black sea bass, with trips containing deeper water species excluded. The commercial line index was based on trips carried out in black sea bass habitat and used a delta lognormal GLM for standardisation. The headboat discard index was generated from all trips in the observer database that discarded black sea bass and standardised with a gamma GLM. Although the time-series is short, it was considered a useful indication of recruitment in recent time.

Indices not selected included: a shallow water trawl index that was rejected on the basis that it was carried out on substrata that were generally not those commonly frequented by black sea bass, a commercial trap logbook index rejected as redundant in the light of fishery-independent indices, potentially not tracking population trends and not used previously, a South Carolina charterboat index which was limited in spatial coverage and not necessarily considered representative, and Florida pot logbooks which had limited spatial and fleet coverage.

Abundance indices are given high weight in the assessment, and the DW described their derivation in detail. Emphasis was rightly given to fishery-independent indices, those verified by observer data, and those with good spatial, temporal and fleet coverage. The selection made provided good temporal coverage of the assessment period with fishery-dependent and -independent indices available (and in broad agreement) for most of the time-series. Possible issues with the reliability of the headboat data, particularly during the early part of the series, were addressed through adjusting the CVs (weights) for this index and alternative assessment runs in the sensitivity analysis.

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

The primary stock assessment model was the Beaufort Assessment Model (BAM), a forward-projecting statistical catch-at-age model implemented in AD Model Builder. The model also accepts and fits to size-structured data, by modelling distributions of size at age. This model has been widely applied to stocks in the region and was used previously for assessment of black sea bass, although the configuration in this assessment has changed substantially. The model fits landings, discards and abundance indices using lognormal likelihoods, and age and length compositions are fitted with multinomial likelihoods. In addition, a lognormal likelihood was applied to the stock–recruitment relationship. Likelihood components are fitted by iterative reweighting. User-supplied weights are applied to landings and discards to achieve a close fit, but allow a little imprecision. Numbers of trips were used as effective sample size for multinomial components, and these initial weights were adjusted until standard deviations of normalized residuals approximated 1.0. Weights on four of the indices (excluding headboats) were then increased upwards to a value of 2.5, to give these indices primacy in the assessment.

Penalty functions were used to maintain estimated parameters within reasonable parameter space, and input variables were sampled from prior distributions to maintain reasonable values.

Fishing mortality was modelled separably, with an overall fishing mortality rate (F) and selectivity fixed by fleet and for time-periods corresponding to size-limit regulations. Selectivities for fisheries landings and discards were modelled parametrically, with landings having flat-topped logistic curves and discards having dome-shaped curves based around a normal distribution of size at age for fish above the minimum size limit. In the past two years when quota restrictions were in place, discard selectivities were modelled as a combination of selectivities of sublegal fish and landed fish weighted by means of fleet-specific observed discards or landings. There was some discussion by the panel regarding the potential for dome-shaped selection curves, but it was felt that the distribution of black sea bass was relatively shallow and that there was not strong evidence for older/larger fish moving into deeper water and becoming unavailable, or for other mechanisms such as direct reduction in gear selection of larger fish or size-structured movements associated with spawning. Therefore, fish were considered equally available to the fishery, and flat-topped selectivity was deemed appropriate.

A biomass dynamics model (ASPIC) was also run to provide an alternative assessment with different population dynamics assumptions.

Data used for the models were discussed under ToR 1, and appropriate measures were taken to bridge gaps where these were incomplete and to address issues regarding their reliability through the assessment weighting functions. Spawning stock was modelled as the annual fecundity of mature females measured at the time of spawning. This took account of the protogynous nature of black sea bass, and discussions about the potential for sperm limitation suggested that this was a problem. Nevertheless, it would be worth outputting a time-series for mature male biomass in future assessments so that the status of this population component can be monitored.

The assessment models used were appropriate for the assessment given the data available. BAM makes use of length and age composition data and provides more detail regarding the population dynamics (e.g. stock–recruitment relationship), but it is configured to fit strongly to the catch and abundance indices data. It is a sensible choice as primary model. Using an alternative simpler model (ASPIC) provides a useful check, although the configuration of the BAM is such that the same data are driving the two models.

Diagnostics from BAM suggested that, in general, the model fitted reasonably well. However, the length data exhibited a number of systematic patterns, with peaks in catch at size at or just above 20cm tending to be underestimated for the MARMAP blackfish traps and the headboat data most noticeably through the 1980s and early 1990s. The headboat data also tended to overestimate catches of fish <20 cm from 1993 on, and the MRIP length data generally fitted poorly. Headboat discard length compositions fitted very closely. Age compositions generally appeared to fit well for all fleets. Abundance indices generally fitted well with a few exceptions for some years in some fleets. Decline in abundance was fitted closely by both MARMAP indices during the 1980s and early 1990s, by the initial points of the commercial line index and by the headboat index. Recent increases in abundance were suggested by the commercial line, headboat and headboat discard indices, and possibly by the most recent point in the MARMAP chevron index.

The indices were positively correlated, indicating significant agreement between them, and provide a sound basis with which to estimate underlying stock trends. Both the BAM and ASPIC therefore provide qualitatively consistent characterisations of stock development.

The fit to the stock–recruitment relationship gave a slight indication of reduced variance and potentially small bias in recent years. This is discussed further in the section on projections (ToR 6).

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

The BAM base run was considered appropriate as the basis for stock assessment for black sea bass. Base-run results were relatively central to the MCB results, with slightly asymmetric 5th and 95th percentiles for SSB relative to reference points. Sensitivity runs suggested that the results from the base run may be slightly optimistic, and there was little evidence for retrospective bias. No direct comparisons of absolute values for biomass and exploitation rate between BAM and ASPIC were made, but SSB and fishing mortality relative to MSY reference levels were qualitatively similar.

Point estimates from the terminal assessment year were

- $F_{2010}$ = 0.702
- $B_{2010}$ = 3796 t
- $SSB_{2010}$ = 1.73E12 eggs

4. Evaluate the methods used to estimate population benchmarks and management parameters (e.g., MSY, $F_{MSY}$, $B_{MSY}$, MSST, MFMT, or their proxies); recommend appropriate management benchmarks, provide estimated values for management benchmarks, and provide declarations of stock status.

The BAM base run provides an appropriate basis for deriving management benchmarks, and relative stock status was broadly consistent between the age-based and surplus production models. Sensitivity analyses suggested that the base model runs may be slightly optimistic, but most runs were consistent with regard to stock status relative to MSY reference points. MCB uncertainty analysis was consistent in indicating with high probability that $SSB<SSB_{MSY}$, while the probability that F exceeds $F_{MSY}$ is only marginally above 50%.

Point estimates for management benchmarks were estimated as:

- MSY = 1.767M lb whole weight
- $F_{MSY}$ = 0.698
- $B_{MSY}$ = 5399 t = 11.9M lb whole weight
- $SSB_{MSY}$ = 2.48E12 eggs
- MSST = 1.54E12 eggs

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).

The method of projecting a randomly selected subset of the MCB runs provides a good basis for projecting the population forward with uncertainty while retaining some of the covariances between variables in the population. Projections implemented the stock–recruitment relationship with lognormal bias correction such that they were consistent with the deterministic projections used to derive management reference points. The review panel requested additional long-term projection runs to check that this was indeed the case, and these were completed and presented, showing that the projected means did correspond to the deterministic values (Table 1.)

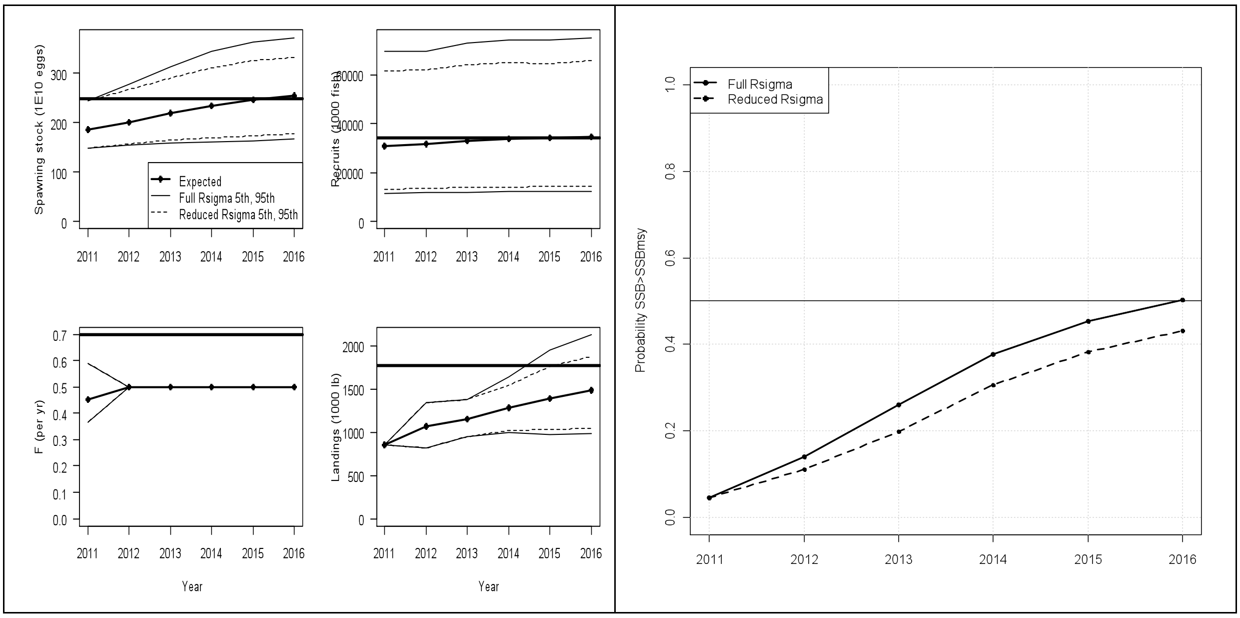
Table 1. Expected values and equilibria from long term projections with $F = F_{MSY}$

| Quantity | Expected | Projected mean* | Projected median* |
|---|---|---|---|
| SSB (1E10 eggs) | 248 | 245 | 236 |
| Landings (1000 lb) | 1767 | 1718 | 1662 |
| Recruits (1000) | 34,393 | 34,892 | 29,912 |

*Means and medians are taken across replicates within years, and then across last 20 years

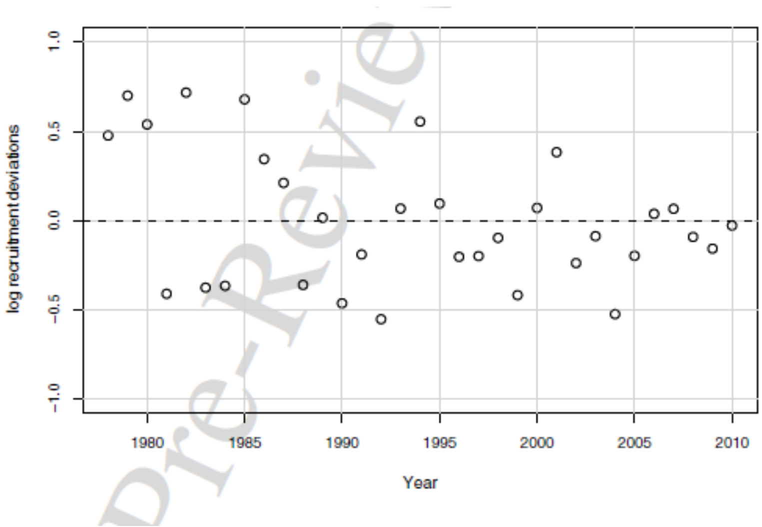
The RP discussed the fit of the stock–recruitment model in the base run, noting that there was a reduction in variance and possibly slight bias in recent years, although consensus was not reached on the latter. Both these observations may well be related to the fact that the high recruitments and biomasses are early in the time-series and that both biomass and recruitment have been low in recent years. It may be that the error structure is not quite capturing differences over time, or it may reflect changes in sampling error through time. The review panel requested further projections to explore the implications of reduced recruitment variation in recent years, which were provided and from which Figure 1 is extracted. The review panel felt it important to stress that these should be exploratory only and not used as the basis for management. Figure 1 shows that under the current rebuilding strategy, reduced recruitment variability would mean that there was a slightly reduced (43% rather than 50%) probability of reaching the existing $SSB_{MSY}$ target, but this would likely be reduced downwards if the change in recruitment was considered to be a real effect. Other projections carried out with the rebuilding strategy modified to meet the existing $SSB_{MSY}$ target under reduced recruitment variability conditions suggested a cost in terms of landings of just under 10%.

Figure 1. Projection using $F = F_{rebuild}$ as described in the assessment report, with 2011 landings at 100% of quota (Scenario 1)




The review panel noted that the effects were relatively small, and consensus regarding the bias was not reached. In my view there is a small bias, with 10–11 negative residuals (of which 1–4 are small) and 4 positive residuals (of which 2 are small) since 1995. These suggest that recent recruitments have been lower than expectations from the stock–recruitment curve and that projections will over-predict recruitment. However, the bias is relatively small, and it is not clear whether it is a real effect or an artefact. If recruitment residuals are available from the MCB runs, then these could be checked relatively straightforwardly (e.g. by taking the mean/sum of residuals since 1995) to see if this or a similar bias occurs in recent years in a large proportion of these runs. If this is the case, then it suggests that the effect is attributable to some signal in the data, although it could still be an artefact of changes in data quality. If the effect is data-driven, a consideration of the data causing it would be advisable, and a small meta-analysis could also be considered to see if there are similar effects in similar stocks. The section on fit diagnostics (ToR 2) noted some anomalies in the fits to some of the length compositions that showed different patterns of over- and underestimation pre- and post- the early 1990s.

Figure 2. Estimated log recruitment residuals. Reproduced from Figure 3.17 of the black sea bass assessment report.




6. Evaluate the adequacy, appropriateness, and application of methods used to characterize uncertainty in estimated parameters. Provide measures of uncertainty for estimated parameters. Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty. Ensure that the implications of uncertainty in technical conclusions are clearly stated.

Uncertainty was explored through a combination of methods including a Monte Carlo parametric bootstrap (MCB), sensitivity runs, retrospective analysis and an alternative assessment assuming different population dynamics. These provide a sound basis with which to characterise and quantify uncertainty in the estimated parameters. The AW report presented the results and implications clearly.

Sensitivity runs investigated alternative BAM model configurations, including alternative input values for M, steepness, model weightings, catchability increasing through time, a continuity run and two runs restricting and removing the headboat index. Industry working documents had expressed concern regarding the headboat data, but the two sensitivity runs either truncating it or removing it completely resulted in very little departure from the base run (in terms of terminal status estimates).

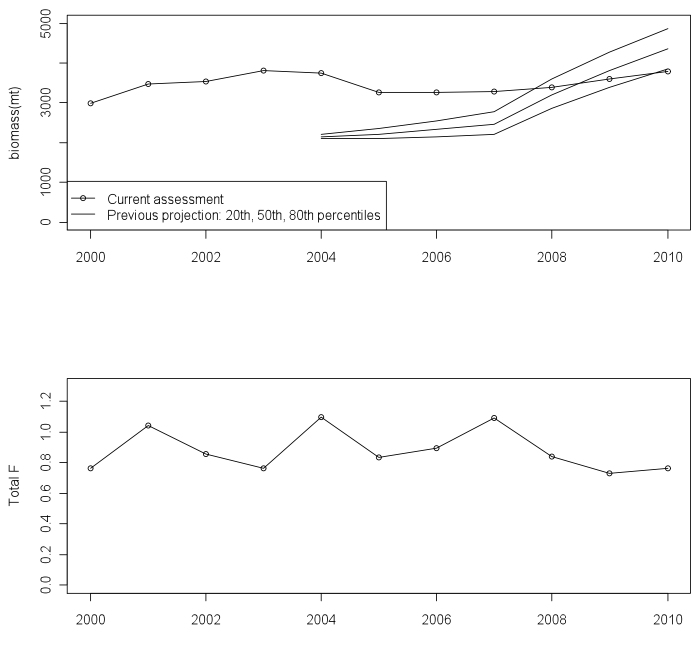
Three of the sensitivity runs were retrospective analyses sequentially removing data back to 2008. Retrospective analyses were limited because three of the input datasets are short time-series available only in recent years, but gave little indication of retrospective bias.

MCB estimated probability density modes, for MSY related benchmarks were consistent with deterministic values for $F_{MSY}$, but slightly below for MSY and well below for $SSB_{MSY}$. Modes of probability densities for current status relative to MSY reference points were consistent with deterministic values for SSB/MSST and $SSB/SSB_{MSY}$, but suggested that current F might be just under $F_{MSY}$, and that the deterministic estimate was just above $F_{MSY}$. Phase plots of output ($SSB/SSB_{MSY}$ vs. $F/F_{MSY}$) from the MCB of the base run showed that the point estimate was very slightly off centre of the distribution and was consistent with the sensitivity perspectives (overfishing and not rebuilt to $SSB_{MSY}$). However, much of the distribution from the MCB lay in the $F<F_{MSY}$ region (see comment on probability density above) and a small proportion suggested $SSB>=SSB_{MSY}$.

The review panel requested that the current assessment results be compared with a projection from the previous assessment (SEDAR 2). This was provided (Figure 3), although it was pointed out that it had been difficult to identify the projection corresponding to the current management plan and that many input parameters had been changed for the current assessment (including M, fecundity, discard estimates, age composition data, model component weights) and that this may have resulted in a rescaling of the assessment outputs. The historical projection also incorporated a lower level of variability. The historical projection indicated stock rebuilding, while there was little evidence of any significant recent increase in SSB in the current assessment or any significant decline in fishing mortality. Discussion noted some overshoot (overage) of quota, which looks likely to transpire again this year.

Alternative model dynamics provided by ASPIC gave qualitative agreement with BAM. Results therefore appeared consistent given available data on uncertainty, across a range of BAM model configurations and for different population dynamics assumptions.

Figure 3. Comparison of biomass from a previous projection with that estimated by the current assessment




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.

Stock assessment results were clearly presented in the AW report and the results are consistent with recommendations made by the review panel. The report was well structured and the statements at the start of the AW stating whether, where and how the ToRs had been addressed were useful.

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

The SEDAR process as applied to the black sea bass assessment was really effective. The reports from both the DW and AW were comprehensive and provided an in-depth consideration of the data, assessment methodology and results along with pertinent assumptions and issues stated. Insufficient time was available for the DW to provide maps showing the distributions of catch and effort for commercial and recreational fisheries, which would have provided useful background, but all other ToRs were met by the DW and AW.

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.

Several research proposals were advanced by the data and assessment workshops, and these were considered in general terms by the review panel with a view to prioritising them without commenting on detail.

In terms of black sea bass, where the recreational sector contributes a substantial part of the catch, improvements to programmes monitoring the recreational sector would have a good impact. However, the introduction of any changes should be accompanied by studies to quantify the effect of these on the time-series, so that the integrity can be maintained and, if required, efforts made to calibrate historical data.

Other priorities for black sea bass would include further studies into the life history and in particular sex transition. The assessment currently uses a single logistic curve to characterize sex transition for the whole time-period, but it is very likely that this may change in response to population demographics or environmental effects (e.g. temperature). Further study into the number of batches spawned per year would also be useful, because this process may be related to fish size or environmental conditions. As both these processes feed into the calculation of spawning potential, they have a critical impact in determining one of the output metrics for stock status.

As the main assessment model is age-structured, improved depth and coverage of age sampling would be beneficial. Similarly, improvements to fishery-independent indices would directly benefit assessment. This could include improvement to the spatial coverage of the existing MARMAP chevron survey or development of a new index.

Spatial structuring was raised as an issue, and studies have been carried out that support current stock identity and management areas, so better spatial information relating to the fishery and suitable habitats for black sea bass may be useful in highlighting areas of potential localized depletion or where spatial management approaches could be successful.

Compilation of historical foreign landings is not considered a priority for this species.

The time-scale for assessment was discussed by the review panel, which considered it primarily a management decision. It needs to take account of the rate at which changes in the stock and fishery can occur (dependent on population and fishery dynamics) and the current status of the stock. In the case of black sea bass, where the stock is subject to a rebuilding plan and has a relatively short lifespan, assessments will be required on a relatively short time-scale, to monitor progress against the rebuilding plan. The current assessment provides a basis for management, so a full benchmark assessment may not be required immediately, but an update within the time-frame of the rebuilding plan may be warranted, especially if additional management measures (e.g. fishery closures) continue to be applied.

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. Complete and submit the Peer Review Summary Report no later than October 28, 2011.

A peer review summary was prepared by the review panel within the specified deadline.

## Golden Tilefish

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

The data workshop listed 24 working documents, six of which related to golden tilefish although none were explicitly reviewed. They included papers on development of standardised catch per unit effort indices, length frequency data, by-catch, discards and other issues in the fishery. The range of documents submitted to the DW provided a good basis for consideration of the data issues.

***Spatial factors***

Evidence to confirm the stock identity for golden tilefish is scarce. A genetic study supported separation of the mid- and south Atlantic stocks, but did not sample North Carolina or southern Virginia, so the boundary between stocks was imprecise. This study found no definitive separation for Gulf of Mexico and South Atlantic stocks, but geographic boundaries are believed to keep the stocks separate. Given the available evidence, the stock and management area defined for golden tilefish is appropriate for stock assessment, but further work into stock structure for the species is required.

Golden tilefish have specific habitat requirements, and submersible and tagging studies suggest that adult fish rarely move more than 2 km. This relatively sedentary habit means that they are potentially more vulnerable to local depletion and that dynamic pool assumptions made in stock assessment models (e.g. all fish equally available to the gear, instantaneous mixing) are more likely to be violated.

***Life history parameters***

Dimorphism was observed in golden tilefish growth, males attaining larger size than females, although the range of sizes and ages was similar for both sexes. As golden tilefish are gutted at sea, sex information is not available for the catch, and a single growth curve was therefore compiled for combined sexes. Although no minimum size limit exists for golden tilefish, the growth curve was adjusted to account for the fastest growers entering the fishery, by assuming 290 mm total length, TL, as an arbitrary minimum size limit and applying the Diaz *et al.* (2004) correction.

A wide range of estimates for natural mortality were calculated for separate and combined sexes, using different mortality formulations and growth parameters. The $\text{Hoenig}_{\text{fish}}$ point estimate of 0.1 was proposed, a decrease relative to the SEDAR 4 data workshop recommendation. Natural mortality was assumed to be size-/age-structured following the Lorenzen (2005) approach. Age-dependent M was higher in this assessment than in SEDAR 4, consistent with a reduction in the perception of longevity following a change in age-determination methodology.

The DW noted that some investigators had suggested that golden tilefish may be protogynous hermaphrodites owing to differences in size distribution by sex and the presence of female tissue in male gonads. However, other authors disagree, pointing to very low levels of previtellogenic oocytes in testes and no other structural evidence for hermaphroditism, and suggesting that the skew in sex ratios by size is attributable to dimorphic growth. The latter opinion was recommended for this stock assessment. Tilefish are batch-spawners, with spawning peaking from April to June. No estimates of annual fecundity at age were available; a relationship for fecundity with size was available, but this was for the total oocytes in the ovary, rather than batch size, and it did not yield an estimate of annual fecundity. SEDAR 4 had provided some information, but recent analyses suggested that this was based on tissue subsamples too small to be reliable. The DW recommended use of a relationship of gonad weight to fish weight or length. Maturity at age or length could not be successfully modelled for separate sexes, and the DW recommended an estimate of $L_{50}$% at age 3, with proportions at ages based on the schedule used for the Gulf of Mexico stock, but shifted upwards by one age. The DW recommended using a sex ratio of 1:1, which generally applied to ages 2–17. Older fish tended to have a greater prevalence of females, though, and there was also some evidence for spatial segregation of the sexes, with more females in fishery-dependent samples from deep water than were noted in fishery-independent samples from the same area, but shallower.

The biology of the golden tilefish is not well understood or quantified, and there are uncertainties regarding many of the life-history parameters required for a detailed analytical stock assessment. The DW summarised the extent of current knowledge and made sensible recommendations with regards to defining biological parameters. However, there is clearly a need for further studies into the biology of golden tilefish if this species is to be assessed routinely. Better data are required across the full range of life history, with a focus on clarifying the reproductive cycle as a priority.

***Discard mortality***

Tilefish are vulnerable to barotraumas, and 100% mortality of discards has been assumed in stock assessments where discards are modelled. However, bycatch and discarding of golden tilefish is low overall (no minimum size limit and an offshore distribution) and it was decided not to model discarding in the assessment.

***Catch***

The golden tilefish fishery is primarily a commercial longline fleet along with commercial handlines. A small proportion of landings is taken recreationally, and some trawling was also noted. The DW discussed misidentified and unclassified landings of tilefish, and concurred with the SEDAR 4 decision to treat goldface tilefish landings as golden tilefish. The pro-rating of unclassified tilefish landings into blueline and golden tilefish carried out at SEDAR

4 was also considered to have been appropriate. All available data were provided, a time-series from 1950 to 2010. The DW compiled landings by state, with rational decisions made for the management area boundary dividing Florida. The final landings by State were derived:

NC: 1958–1993 (ACCSP), 1994–2010 (NC DMF)
SC: 1958–1979 (ACCSP), 1980–2010 (SC DNR)
GA: 1958–2010 (ACCSP)
FL: 1958–1985 (ACCSP), 1986–2010 (FL trip ticket)

Recreational landing were compiled from the MRFSS and SRHS. The AW modified the recreational landings slightly by substituting the 2005 value with the mean of the adjacent four years.

Low sample size precluded the calculation of discards. Fishers report that they are able to avoid tilefish habitat during closed seasons, so bycatch is considered low and there is no minimum size limit. The DW concluded that discards are probably few in number and unlikely to affect the assessment.

The DW accepted many of the earlier SEDAR 4 recommendations in compiling landings data in a consistent and reasoned manner. The decision not to include discards in the assessment is also appropriate given the low prevalence of discarding and indeed the paucity of data.

***Age and length distributions***

Age and length data were obtained from the NMFS TIP programme with quality control on record selection from the database. Sufficient trips were sampled for length data in most years (not 1987–1990) for commercial longline, but only in 2002 for handline, whereas other gears were rarely sampled. With some exceptions, age compositions were available for commercial handlines and longlines for the years 1984–2010.

Age determination of tilefish is not straightforward and the AW introduced an ageing error matrix to the assessment based on the SEDAR 25 Gulf of Mexico tilefish assessment. Several levels of adjustment were evaluated, with the final ageing error halved to improve the fits to age compositions.

Decisions made by the DW and AW were soundly based with the intention of maximising data content available for assessment. Golden tilefish is data-poor in many respects and further work to improve age determination and quantify associated errors would be beneficial.

***Abundance indices***

Three abundance indices were considered, one fishery-independent and two commercial. MARMAP longline (1983–2010) was selected as a fishery-independent index although it had small sample size and high variance, was not standardised by GLM, and because of small sample size in many years, CPUE was calculated for four-year blocks rather than annually.

Commercial longline (1993–2010) was selected as a fishery-dependent index. It is based on logbook data and calculated as weight per hook fished because soak time was not consistently reported. Trips were selected using an algorithm that analysed species compositions and minimised both predictions of tilefish presence when they were absent (false positives) and tilefish absence when they were present (false negatives). This index was selected for use in the assessment.

A short time-series provided by industry representatives and based on two vessels was not used because it should already contribute to the commercial longline index.

Decisions made with regard to abundance indices made best use of the available data by selecting a fishery-dependent index based on the dominant gear and modifying a fishery-independent index to provide a general signal over a multiyear period so that it could be included.

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

The primary assessment model was BAM. The general assumptions are described in the AW report and commented on in this review under black sea bass above. In general terms this model is considered an appropriate choice for making best use of the data available (i.e. landings, abundance indices, age and length compositions) and provides a good basis for projection and uncertainty analysis.

The configuration was appropriate for the biology of, and data available for, golden tilefish. Details specific to the golden tilefish assessment included:

i) upweighting of abundance indices was applied only to the commercial longline index (weight = 3.0) and not to the MARMAP line index.
ii) Steepness for the Beverton & Holt stock–recruitment relationship could not be estimated and was fixed at a value of 0.84 (MC'd for MCB analysis) on the basis of a meta-analysis.

The AW accepted that this configuration may not be the best representation of reality, but considered that it provided a basis for assessment and projection and that its suitability would be evaluated by sensitivity and uncertainty analyses.

Understanding of golden tilefish biology and population dynamics is limited, and the stock is data-poor. The configuration of the BAM model applied made best use of available data to provide a basis to derive management benchmarks and carry out projections. MCB, sensitivity analyses (including retrospective runs) and an alternative assessment model help to clarify the uncertainty associated with the assessment.

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

Terminal assessment estimates from BAM were:

$F_{2010} = 0.075$
$B_{2011} = 5244$ t
$SSB_{2010} = 54.8$ t gonad weight

It should be noted that outputs from BAM and ASPIC, although qualitatively similar, were scaled quite differently. In both cases, model dynamics depended on the functional form of the model (stock–recruitment model vs. production model) and assumptions regarding the low level of exploitation early in the time-series. The absolute values of these output metrics should therefore be treated with some caution.

4. Evaluate the methods used to estimate population benchmarks and management parameters (e.g., MSY, $F_{MSY}$, $B_{MSY}$, MSST, MFMT, or their proxies); recommend appropriate management benchmarks, provide estimated values for management benchmarks, and provide declarations of stock status.

Methods used to estimate management benchmarks were based on current best practice and consistently implemented. However, as with estimates of stock abundance, biomass and exploitation, the scale is dependent upon model type and assumptions, and they need to be considered and applied in a (model) consistent manner rather than in absolute terms.

Further, MSY benchmarks depend to a large extent on the stock–recruitment relationship, for which in this case steepness could not be estimated and was fixed on the basis of a meta-analysis. Although the AW explored a range of alternatives to try and estimate steepness, they were unsuccessful. The fitted stock–recruitment relationship included a single very large residual for the 2000 year class, which may result from problems with age determination in this species, and the MCB analysis indicated that the $R_0$ parameter was substantially biased from the mode. Increasing penalties to improve the stock–recruitment relationship fit resulted in a poor fit to the main commercial longline abundance index. There is therefore a considerable degree of uncertainty associated with the productivity of this stock.

Management benchmarks estimated using BAM were:

MSY =638k lb whole weight
$F_{MSY}$ = 0.185
$B_{MSY}$ = 2918 t = 6.4M lb whole weight
$SSB_{MSY}$ = 25.3 t gonad weight
MSST=22.6 t gonad weight

Comparison with ASPIC showed that stock status ($F/F_{MSY}$ and $B/B_{MSY}$) was broadly similar between the catch-at-age and production models despite differences in absolute values.

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).

The method of projecting a randomly selected subset of the MCB runs provides a good basis for projecting the population forward with uncertainty while retaining some of the covariances between variables in the population. Projections implemented the stock–recruitment relationship by bootstrapping recruitment residuals because this was felt to represent the uncertainty in the fit better. The fixed steepness parameter was derived by Monte Carlo sampling around the point value. This approach provides a method for projection that is consistent with the assessment and effectively incorporates uncertainty.

Projections were carried out for a range of fixed F scenarios ($F_{MSY}$, $F_{current}$, 65%$F_{MSY}$, 75%$F_{MSY}$ and 85%$F_{MSY}$), all of which maintained SSB above $SSB_{MSY}$ with 50% or more probability. Fishing at $F_{current}$ resulted in SSB well above $SSB_{MSY}$ with high probability, and provided expected yields slightly lower than MSY.

6. Evaluate the adequacy, appropriateness, and application of methods used to characterize uncertainty in estimated parameters. Provide measures of uncertainty for estimated parameters. Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty. Ensure that the implications of uncertainty in technical conclusions are clearly stated.

Uncertainty was explored through a combination of methods including a Monte Carlo parametric bootstrap (MCB), sensitivity runs, retrospective analysis and an alternative assessment assuming different population dynamics. These provide a sound basis with which to characterise and quantify uncertainty in the estimated parameters. The AW report presented the results and implications clearly.

The MCB analysis phase plot indicated that status as estimated by the base run was optimistic, falling towards the lower margin of the distribution of $F/F_{MSY}$ and above the main cluster of points for SSB/MSST. This is confirmed by the probability distribution plots for these status indicators. However, the vast majority of points indicate that SSB > MSST and $F < F_{MSY}$. $5^{th}$ and $95^{th}$ percentiles of the distribution show that there is a >5% chance for both F exceeding $F_{MSY}$ and SSB being below MSST. The MCB was considered sufficient to provide the basis for a P* analysis for setting future management metrics.

In all, 24 sensitivity runs were carried out including eight retrospective runs. Sensitivity runs concurred with the MCB that the base run may be optimistic, but all runs indicated SSB > MSST and $F < F_{MSY}$. Retrospective analyses did not indicate strong systematic bias in output metrics, although SSB and biomass estimates for 2005–2007 have been revised upwards in recent assessments.

BAM and ASPIC produced qualitatively similar results for stock and fishery status. A comparison of productivity as determined by BAM and ASPIC was requested and additional analyses were carried out and presented (Figures 4 and 5). Both are to a large extent predicated by assumptions regarding initial biomasses, set close to unfished levels in both cases, and functional forms of models, with ASPIC having a fixed functional form (logistic), while the BAM is driven by the Beverton & Holt stock–recruitment function with fixed steepness. Although they produce different absolute outputs, they are similar in terms of status relative to MSY reference points.

Figure 4. Comparison of BAM and ASPIC annual estimates of $F/F_{MSY}$ for golden tilefish

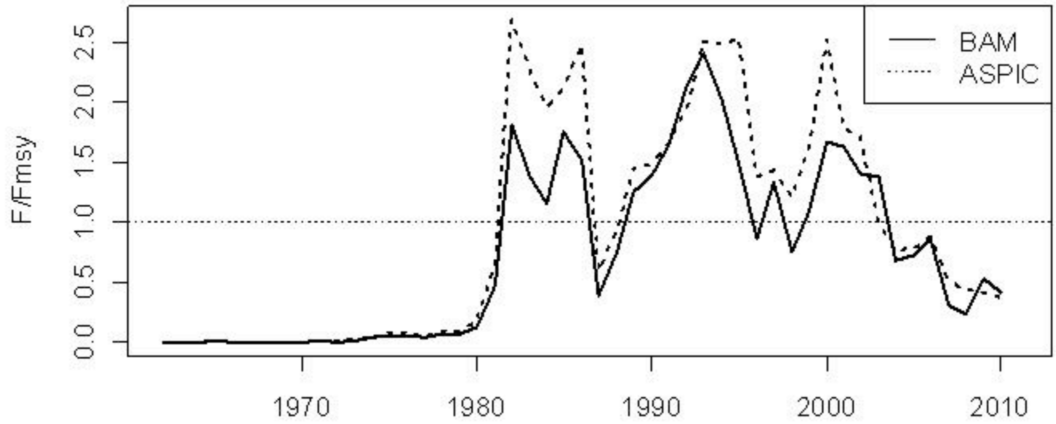
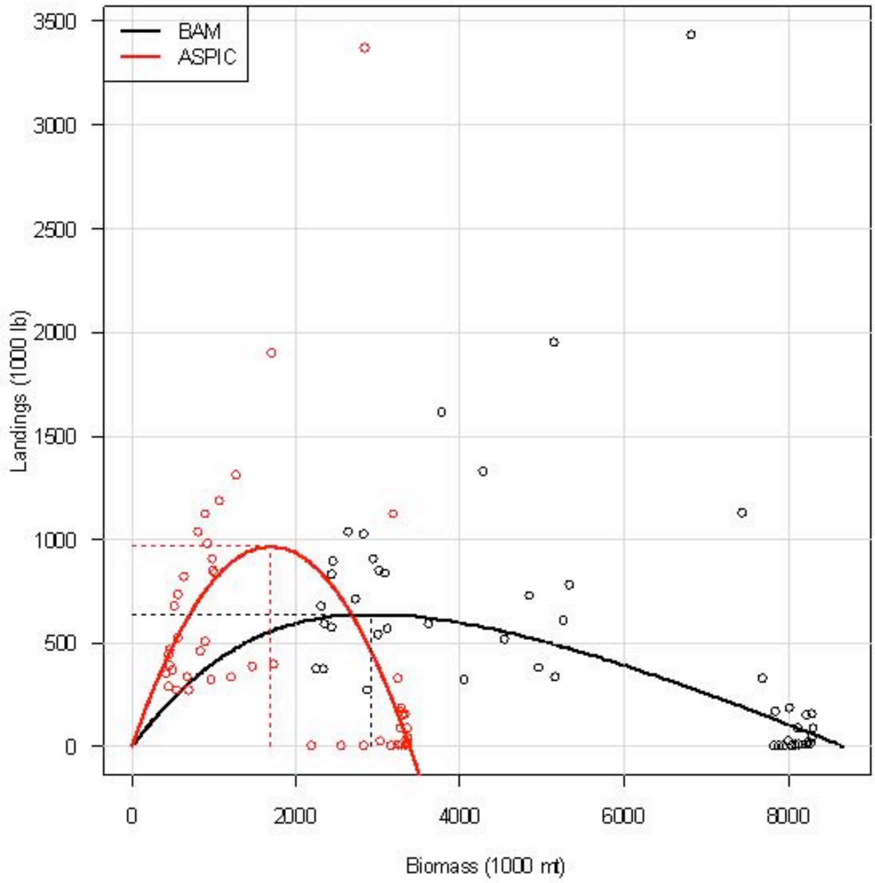



Figure 5. Comparison of BAM and ASPIC models for golden tilefish




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.

The stock assessment results were clearly and concisely expressed by the AW report and consistent with recommendations from previous reports.

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

The SEDAR peer review process provides a rigorous standard and transparent review process. The DW addressed its terms of reference with the exception of providing maps of catch and effort distribution, for which there was insufficient time. The AW addressed all its terms of reference.

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.

Studies to improve quantification of life-history parameters and that improve the quality and extent of length and age data are considered of high priority and likely to improve the reliability of assessment. The development of a reliable fishery-independent index with good coverage would produce similar improvements in future. Improved age and length sampling and survey programmes may also provide benefits for species other than golden tilefish if they are comprehensive.

Improved knowledge of migration and movements would be beneficial to a number of aspects (stock identity, local depletion, dome-shaped selection, etc), but may be a secondary issue.

Improvements to the programmes monitoring the recreational fishery were not considered a high priority for this species, although as with other sampling programmes, if they are well planned and comprehensive, they may be able to provide benefits across a wide range of species.

Compilation of historical foreign landings is not considered a priority for this species.

The timing of future assessments is essentially a management decision, but can be informed by the science. Current stock status is relatively healthy, recent good recruitment is suggested, and the species is relatively long-lived, so new information that will radically alter perceptions is unlikely in the short term and management should be relatively stable. There would therefore seem little utility in conducting either a benchmark or update assessment in the near future. In the medium term, an update may suffice unless new information or management requirements have come to light.

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. Complete and submit the Peer Review Summary Report no later than October 28, 2011.

A peer review summary was prepared by the review panel within the specified deadline.

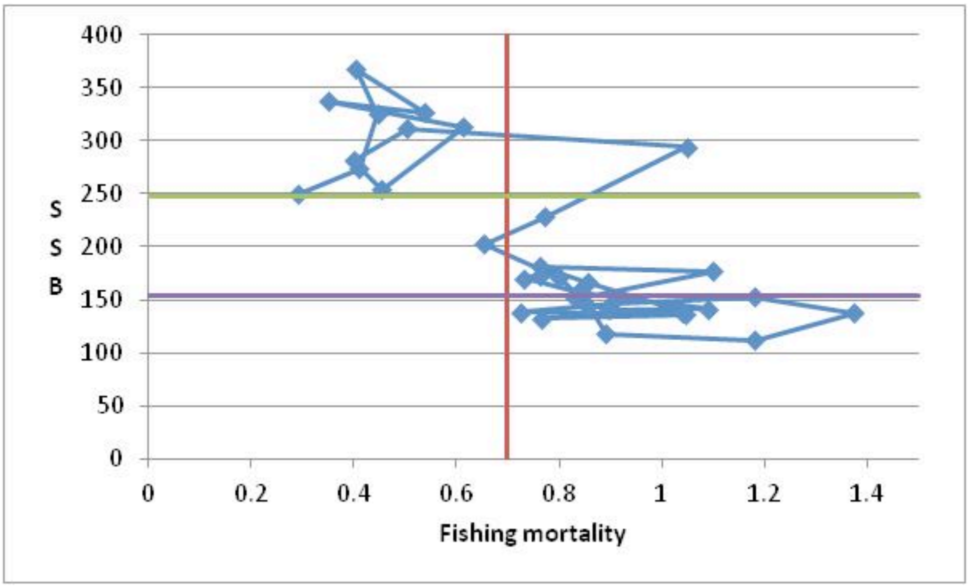
## 5. Conclusions and Recommendations

The black sea bass assessment indicated that F was above $F_{MSY}$, and although SSB has increased above MSST, it has not yet reached $SSB_{MSY}$, the rebuilding target. This result was generally consistent over a range of uncertainty in data, model configurations and population dynamics assumptions. The stock–recruitment relationship caused some concern with temporal structuring apparent in the residuals. Whether these reflected a real effect or an artefact was not clear, but this merits further investigation, initially via analysis of stock–recruitment fits in the MCB analysis, which if available in the assessment outputs could indicate whether or not the effect is data-driven. The major source of uncertainty in the black sea bass assessment was considered to relate to monitoring of the recreational fishery, which is the major component of the landings. Improvements to monitoring programmes for the recreational sector would therefore be beneficial, although maintaining consistency with historical time-series is also crucial.

The golden tilefish assessment indicated that the stock is currently neither overfished nor suffering overfishing. This result was generally consistent across a range of uncertainty in data, model configurations and population dynamics assumptions. The stock–recruitment relationship was one of the major sources of uncertainty because the steepness parameter could not be estimated, and a single large residual also gave some cause for concern. The lack of a reliable fishery-independent index was also considered a significant source of uncertainty. Knowledge and data relating to the golden tilefish stock are lacking, with poor quantification of life-history parameters, relatively low levels of biological sampling, and difficulties in age determination. There is substantial scope for further research on golden tilefish if this can be politically or economically justified.

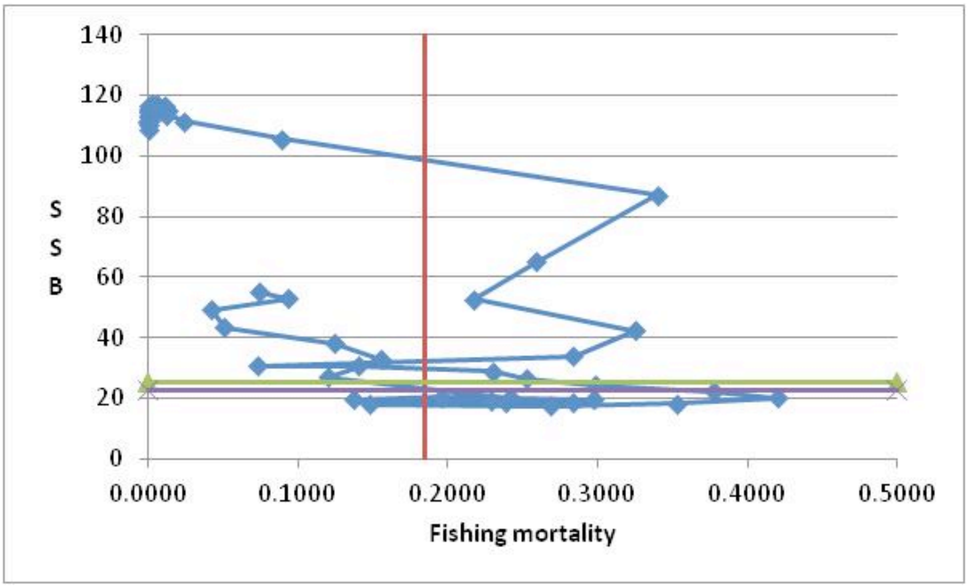
Reviewers were also asked to indicate any analyses or outputs that could enhance the stock assessment report. A graphic used in ICES reports that can be informative is a phase plot of SSB on F showing the trajectory of the stock and fishery through time (e.g. Figures 6 and 7 below, for black sea bass and golden tilefish, respectively). In some cases, these plots show a clockwise trajectory and provide an empirical evaluation of the levels of F at which the stock tends to rebuild or decline. In the current examples, this is indeed the case for golden tilefish, SSB declining sharply with fishing mortality levels >0.25, but the stock rebuilds with fishing mortality below or around 0.15. The plot for black sea bass does not show this pattern, but might be expected to do so if lower F could be achieved.

Figure 6. Phase plot of black sea bass stock and fishing mortality trajectory



Benchmark reference points indicated: $F_{MSY}$ (red), $SSB_{MSY}$ (green) & MSST (purple)

Figure 7. Phase plot of golden tilefish stock and fishing mortality trajectory



Benchmark reference points indicated: $F_{MSY}$ (red), $SSB_{MSY}$ (green) & MSST (purple)

## Appendix I. Material provided for the review

| Document # | Title | Authors |
|---|---|---|
| | **Documents Prepared for the Review Workshop** | |
| SEDAR25-RW01 | Comments and notes received during the data, assessment and review for SEDAR 25 | Multiple authors |
| SEDAR25-RW02 | Comments and notes received during the assessment and review for SEDAR 25 | Multiple authors |
| SEDAR25-RW03 | The Beaufort Assessment Model (BAM) with application to black sea bass: model description, implementation details, and computer code | Sustainable Fisheries Branch, NMFS 2011 |
| SEDAR25-RW04 | The Beaufort Assessment Model (BAM) with application to tilefish: model description, implementation details, and computer code | Sustainable Fisheries Branch, NMFS 2011 |
| SEDAR25-RW05 | Development and diagnostics of the Beaufort assessment model applied to black sea bass | Sustainable Fisheries Branch, NMFS 2011 |
| SEDAR25-RW06 | Development and diagnostics of the Beaufort assessment model applied to tilefish | Sustainable Fisheries Branch, NMFS 2011 |
| SEDAR25-RW07 | Use of MARMAP age compositions in SEDAR 25 – Methods of addressing sub-sampling concerns from SEDAR 2 and SEDAR 17 | Ballenger, Reichert, and Stephen, 2011 |
| SEDAR25-RW08 | Fisheries management actions confound the ability of the Beaufort Assessment Model (BAM) to explain dynamics of the Golden Tilefish fishery off of east Florida | Hull and Barile, 2011 |
| SEDAR25-RW09 | A note on the use of flat-topped selectivity curves in SEDAR 25 | Hull and Hester, 2011 |
| SEDAR25-RW10 | On steepness | Hull and Hester, 2011 |
| SEDAR25-RW11 | Some considerations of area interactions | Hull and Hester, 2011 |
| | **Data workshop reports** | |
| SEDAR25-DWR1 | South Atlantic Black Sea Bass, Section II: Data Workshop Report | SEDAR 25, June 2011 |

| | | |
|---|---|---|
| SEDAR25-DWR2 | South Atlantic Golden Tilefish, Section II: Data Workshop Report | SEDAR 25, June 2011 |
| **Final Assessment Reports** | | |
| SEDAR25-SAR1 | Assessment of Black Sea Bass in the US South Atlantic | SEDAR 25 |
| SEDAR25- SAR2 | Assessment of Golden Tilefish in the US South Atlantic | SEDAR 25 |

**Many other background documents and working documents submitted to the data and assessment workshops were also available on the SEDAR website.**

## Appendix 2. Statement of work for Michael Smith (Cefas)

**External Independent Peer Review by the Center for Independent Experts**

**SEDAR 25 South Atlantic Black Sea Bass and Golden Tilefish Review**

**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 25 will be a compilation of data, a benchmark assessment of the stock, and an assessment review conducted for South Atlantic Black Sea Bass and Golden Tilefish. 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 workshop 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 25 are within the jurisdiction of the South Atlantic Fisheries Management Council and the states of Florida, Georgia, South Carolina, and North Carolina. 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 and recent experience in the application 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. Expertise with data poor assessment methods would be preferable. 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 Charleston, South Carolina during October 11-13, 2011.

**Statement of Tasks:** Each CIE reviewer 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/sponsor.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 cannot 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 in the panel review meeting in Charleston, South Carolina during October 11-13, 2011.
3) In Charleston, South Carolina during October 11-13, 2011 as specified herein, conduct an independent peer review in accordance with the ToRs (**Annex 2**).
4) No later than October 27, 2011, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to 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.

| | |
|---|---|
| September 6, 2011 | CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact |
| September 27, 2011 | NMFS Project Contact sends the CIE Reviewers the pre-review documents |
| **October 11-13, 2011** | Each reviewer participates and conducts an independent peer review during the panel review meeting |
| October 27, 2011 | CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator |
| November 10, 2011 | CIE submits CIE independent peer review reports to the COTR |
| November 17, 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, Program Manager, 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:**

Kari Fenske, SEDAR Coordinator
4055 Faber Place Drive, Suite 201
North Charleston, SC 29405
kari.fenske@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 25 Black Sea Bass Review Workshop Terms of Reference

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, Fmsy, Bmsy, MSST, MFMT, or their proxies*); recommend appropriate management benchmarks, provide estimated values for management benchmarks, and provide 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. Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty. 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 assessment 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. Complete and submit the Peer Review Summary Report no later than October 28, 2011.

* 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.

## SEDAR 25 Golden Tilefish Review Workshop Terms of Reference

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, Fmsy, Bmsy, MSST, MFMT, or their proxies*); recommend appropriate management benchmarks, provide estimated values for management benchmarks, and provide 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. Comment on the degree to which methods used to evaluate uncertainty reflect and capture the significant sources of uncertainty. 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 assessment 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. Complete and submit the Peer Review Summary Report no later than October 28, 2011.

* 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**

## Tentative Agenda

# SEDAR 25 South Atlantic Black Sea Bass and Golden Tilefish Review Workshop

Charleston, SC
11-13 October, 2011

***Tuesday***

| | | |
|---|---|---|
| **9:00 a.m.** | **Convene** | |
| **9:00 – 9:30 a.m.** | **Introductions and Opening Remarks**<br>**Coordinator**<br>- *Agenda Review, TOR, Task Assignments* | |
| **9:30 – 12:00 p.m.** | **Assessment Presentations and discussion** | |
| **12:00 – 1:15 p.m.** | **Lunch Break** | |
| **1:15 – 6:00 p.m.** | **Assessment presentations and discussion** | **Chair** |

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

***Wednesday***

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

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

***Thursday***

| | | |
|---|---|---|
| **8:00 a.m. – 1:00 p.m.** | **Panel Discussion**<br>- *Final sensitivities reviewed.*<br>- *Projections reviewed.*<br>- *Review Reports* | **Chair** |
| **1:00 p.m.** | **ADJOURN** | |

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

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

| Appointee | Function | Affiliation |
| --- | --- | --- |
| Jim Berkson | SSC Rep | SAFMC SSC |
| Steve Cadrin | SSC Rep | SAFMC SSC |
| Anne Lange | Review Panel Chair | SAFMC SSC |
| Mike Bell | CIE Reviewer | CIE |
| Paul Medley | CIE Reviewer | CIE |
| Mike Smith | CIE Reviewer | CIE |