SEDAR 28:

Gulf of Mexico

Cobia and Spanish Mackerel

Stock Assessment Review

P.L. Cordue

Fisheries Consultant

New Zealand

***For* CIE Independent System for Peer Review**

19 February 2013

Executive summary

A desktop review of Gulf of Mexico cobia and Spanish mackerel stock assessments was conducted by three independent CIE reviewers, in January 2013, as part of SEDAR 28. This document presents my findings and recommendations, with regard to the assessments, based on a detailed review of the assessments as described in the Data and Assessment Workshop reports and supporting documents.

The cobia and Spanish mackerel stocks in the Gulf of Mexico were both assessed using the Stock Synthesis package SS3. This is a well-tested package which enables fully-integrated age-structured stock assessments using landings, discards, length, and age data from multiple fisheries.

Both assessments used very similar data sources: landings and discard data from recreational fisheries (the bulk of the landings) and some commercial fisheries; discard estimates from the shrimp fishery (substantial in some years); length and age data as available for each fishery; and standardized CPUE indices.

A simple and typical model structure was used in both assessments. Population in age-structured equilibrium before the start of the fisheries. Year-round fisheries with constant selectivity patterns (with some time-blocking). Constant age-specific natural mortality over time. A single von Bertalanffy growth curve estimated in the model and a Beverton Holt stock-recruitment relationship. Year class strengths (recruitment deviations) estimated for about 20 cohorts.

The assessments have common problems: the CPUE time series used in the assessment runs are not defensible as relative abundance indices; and the length and age data were not appropriately post-stratified or scaled. Primarily because of the lack of defensible abundance indices it would be unsafe to use the assessments to provide management advice.

My main conclusions are:

- Stock structure and fixed life history parameters were adequately considered.
- Landings history, discards, and discard mortalities were adequately determined and considered.
- Composition data were poorly treated at both the Data and Assessment Workshops. There was an absence of appropriate analysis and discussion with regard to post-stratification of the data to deal with inadequate sample sizes within some strata.
- The Index Working Group made very poor recommendations with regard to the time series to use in the stock assessments as relative abundance indices:
    - For cobia, two recreational CPUE time series were recommended but these both had very low proportions of successful trips and spanned a period when fishing regulations had become more restrictive.
    - For Spanish mackerel: a SEAMAP survey was recommended as a recruitment time series, but it caught very few Spanish mackerel each year; a recreational time series was recommended but it had a very low proportion of successful trips; and a commercial index based on catch-per-trip was recommended but it had not been standardized for trip duration or time fished.
- None of the abundance indices used in the stock assessment runs are defensible.

- The model structure used, the choice of runs, and the methods of projection and describing of uncertainty were adequate but could not overcome the flawed data inputs.
- None of the model runs should be used to determine biomass estimates or recommend stock status.

My main recommendations are:

- Top priority should be given to the construction of defensible abundance indices for both cobia and Spanish mackerel from the commercial and recreational data:
    - Talk to some of the participants in the fisheries to get an understanding of how, when, and where, they target cobia and Spanish mackerel (if at all).
    - Perform a full descriptive/exploratory analysis of the data to understand the temporal and spatial variation in the catches and the potential explanatory variables.
    - Identify regional and seasonal fisheries for which fishing effort is "likely" to catch cobia or Spanish mackerel.
    - Perform an analysis to determine if fishing regulations have impacted on the ability of the data to track abundance (time series may have to be split to account for different fishing behaviour caused by regulation changes)
    - Produce standardized CPUE indices for each identified regional/seasonal fishery and consider which if any can defensibly be used as abundance indices.

- If defensible abundance indices can be constructed then assessments can be done as before except:
    - Composition data should be appropriately post-stratified and scaled; sample sizes should be based on the number of trips/landings sampled (not the number of fish measured or aged).
    - Recruitment deviates should only be estimated for cohorts which are well-represented in the composition data (e.g., appear at least three times in the age data).

Background

The South-East, Data, Assessment, Review (SEDAR) process was initiated in 2002 to improve the reliability of fishery stock assessments in the South Atlantic, Gulf of Mexico, and US Caribbean. This review is part of SEDAR 28 and covers the Gulf of Mexico Spanish mackerel and cobia stock assessments.

I am one of three CIE reviewers who performed a desktop review during January 2013. The three reviews are meant to be independent and I have had no contact or discussion with the other two reviewers. This report presents my findings and recommendations in accordance with the Terms of Reference (ToRs) for the review (Appendix 2, annex 2).

Review Activities

The main documents provided for the review were made available in a timely manner through an ftp site. Also, a link was provided to the SEDAR website which contained many workshop, background, and reference documents (Appendix 1).

I noted, that in the original ToRs, it was assumed that a normal review was being conducted and that the reviewers would jointly write a Summary Report. I contacted CIE and they supplied me with amended ToRs which were specific to a desktop review (Appendix 2, annex 2).

The main documents for the review were the Data Workshop and Assessment Workshop reports (Appendix 1). I read these four reports in detail, a number of times, over the period of the review and consulted specific workshop or reference documents as needed. I also searched the Web to obtain information on current and past federal and state recreational fishing regulations for cobia (in particular).

Summary of findings

Cobia and Spanish mackerel were both assessed using the Stock Synthesis package SS3. This is a well-tested package that allows data from a range of sources to be fitted to obtain estimates of population parameters and management quantities. Estimates of uncertainty were obtained by performing sensitivity runs and bootstrapping the main runs.

The two assessments use very similar methods and data sources (estimated catch histories for commercial and recreational fisheries, abundance indices, and length and age data). For this reason they share a number of strengths and weaknesses.

Before considering the specific ToRs for each assessment I will discuss some problems which are common to both assessments.

Obtaining abundance indices from recreational CPUE data

For both assessments standardized CPUE indices were calculated for the headboat survey and for the MRFSS data (although the headboat time series was not used in the mackerel assessment). In each case a delta-lognormal model was used (binomial for success/failure and lognormal for positive catches). This approach was applied to the whole of each dataset with limited or no filtering of records to remove irrelevant effort. As a consequence, the proportions of successful trips (those that caught the species of interest) were very low (mackerel: MRFSS 5%, headboat <5%; cobia: MRFSS <1%, headboat 7%).

These success rates are so low that one would think that it was very unlikely that the CPUE indices could be tracking abundance. The Index Working Groups (IWG) had attempted to filter the data to obtain relevant effort using Stephens and MacCall (2004) and a number of ad hoc approaches. However, they were unable to find a satisfactory subset of the data to use and defaulted to the full data set. (The failure of Stephens and MacCall (2004) is interesting and bears further investigation at a later date – why did the method fail so completely?)

I have no faith in any of these CPUE time series as indices of *relative abundance* because the very low success rates show that most of the effort is irrelevant to cobia and Spanish mackerel. This means that the basic assumption of catch being proportional to effort is violated. The standardization of the indices does not help. To get a defensible abundance index from these data requires that relevant effort is identified – e.g., so that a doubling of effort (in a given "stratum") will result in a doubling of catch – or a doubling of biomass for a given amount of effort will double the catch.

In order to subset these data and identify relevant effort it is necessary to obtain an understanding of the different recreational fisheries that are operating on cobia and Spanish mackerel. This will not be an easy process. It will probably require that additional information on the operation of the fisheries be obtained by interviewing the participants (e.g. headboat skippers). Cobia and Spanish mackerel are probably targeted by recreational fishers in some places at some times during the year (e.g., cobia during a known migration wave). It may be possible to identify vessels which fish in certain areas at certain times and to use their data (positive catches and success/failure in the given areas and times) to obtain defensible abundance indices. Alternatively, it may be that additional information needs to be routinely collected from recreational fishers before any reliable abundance indices can be produced from the recreational fisheries for these species.

Using the positive catches is a possibility, which was explored by the IWG. The concern is that such indices will be hyperstable. However, with sufficient descriptive analysis it may be possible to justify the use of just the positive trips (e.g., showing that there is no shrinkage in the area and the season from which successful trips occur over time).

Changes in recreational fishing regulations

Changes in fishing regulations have to be considered when recreational CPUE data are being analysed for abundance indices.

For cobia, the Data Workshop report contains no information on changes in regulations or the variation in regulations between state and federal waters. This is a serious omission because the federal daily bag limit of 2 per person did not come into effect until August 1990 and in Florida state waters the limit was reduced to 1 per person (with no more than 6 per vessel) on 22 March 2001. The only abundance indices used in the cobia assessment are the headboat and MRFSS time series which both span the period of regulation changes (headboat: 1986-2010; MRFSS: 1981-2010). The implementation of a minimum legal size for cobia in 1984 is mentioned in the Data Workshop report and the potential change in selectivity is modelled in the assessment. In the Assessment Workshop report the imposition of the federal bag limit in 1990 is noted, but only in the discussion of the fit to discard rates. The Florida state regulation is not mentioned in the Assessment Workshop report.

For Spanish mackerel there were numerous changes in bag limits over the period covered by the MRFSS CPUE indices. The fact that there were changes is noted in the reports but no analysis or discussion of the potential effect on catch rates is given. The changes were generally increases in the daily bag limit, so it may be that they are not particularly important in terms of affecting catch rates. However, there should have been an analysis of the data to see if there were effects such as a limiting of catch before the bag limits were increased.

Modelling of year interactions as random effects
The standard approach taken by the Index Working Group when standardizing the commercial and recreational CPUE data was to fit two-way interactions involving year as a random effect. The software will let this be done, but it is inappropriate because year interactions are probably not random (in the sense of random effects, where the values can be considered as random samples from a particular distribution). For example, consider a year-area interaction. If there are very different trends in different areas then this is a sign that there are groups of fish associated with each area which have different abundance trajectories – not a random effect at all (the changes in abundance are correlated within each area and perhaps across areas). Also, it is a sign of a fundamental problem with the CPUE analysis. A valid abundance index can only be obtained in this case if the number of records in each area is a good approximation to the relative abundance across areas (so that the different trends are appropriately weighted). Fitting the year-area interactions as a random effect does not change the mean effects (Venables and Dichmont, 2004) and merely hides the potential problem. This is not to say that mixed models should not be used – there are factors which can be appropriately modelled as random effects (e.g., individual vessel effects).

Scaling of length and age (composition) data
It is important to try to make of the most of whatever composition data are available. These are the data that provide information on growth, selectivity, and year class strength. If they are not properly stratified and scaled then legitimate signals in the data will be obscured.

There should be little debate about how length and age data are scaled. If there was an appropriate sampling design, then this includes the stratification and how to scale the data. For length samples, normally, there is a two-stage scaling procedure: sample scaled to catch or landing; and then the combined samples within a stratum are scaled to the stratum catch (and then combined across strata without any further weighting). For age data, sampled at random, the same scaling procedure applies. For age data, collected to construct an age-length key, the length frequency is first constructed (by appropriate scaling) and then the age-length key(s) is applied to produce the age frequency.

The recommendation of the Data Working Group, for both cobia and mackerel, to scale the age data "using the length frequency" is very worrying. I first heard of this method when reviewing SEDAR 17 and on investigation I found that it was invalid. Simple examples were enough to show that the method did not achieve its stated intent (Cordue 2008). That the same method is still being recommended is very disappointing. They cite a paper which apparently uses the method when estimating growth curves (Chih 2009). It may have some utility in the situation the author considered but the method should not be used to produce age frequencies.

When composition data are sampled in an ad hoc basis (or there are inadequate sample sizes in the original stratification) it is important to post-stratify in such a way that the full (spatial and temporal) extent of the fishery is covered with adequate sample sizes in each stratum (for the years, or groups of years, in which there are adequate data). It is also important to exclude data in years when the coverage is inadequate – it should not just be "thrown in" in the hope that the model can account for non-representative samples (because it cannot).

Using age data as conditional age-at-length

This appears to have become the norm for assessments using SS3. It has advantages and disadvantages. It stops the worry about the double-use of age and length data, where the age data came from a subset of the fish that were measured. Also, it allows non-randomly collected age samples to be used in the assessment in a natural fashion and facilitates the estimation of growth parameters. However, it does not preclude the necessity for a careful analysis of the age data in terms of where samples came from, when they were collected, and how they were collected.

One problem is the timing of the sampling. It is important to consider how fast the fish grow and at what size they are recruited to each fishery. If fish are growing rapidly during the year in which they were sampled then there is the problem that the age proportions at given length change during the year (e.g., sample for age at 20 cms: on 1 February the proportions at age are 70% 1 year old and 30% 2 year old; but on 1 November the expected proportions are 100% 1 year old).

Another issue is that age-proportions at given length can also vary spatially. For example, a recreational fishery in one area may be catching spawning fish, while in another area the same "fishery" (in the model at least) is capturing non-spawning fish. The age-proportions at length will be very different between the two areas. A similar effect could occur because of spatial variation in growth. Yet another issue is the variation in growth between cohorts. At a given time of year, the age-proportions at a given length could be dramatically different for fast and slow growing cohorts. If there is only patchy conditional age-at-length data in the model then fast growing cohorts could be estimated as strong cohorts and slow-growing cohorts as weak cohorts.

Because of all of these issues it is by no means certain that it is best to incorporate age data into SS3 as conditional age-at-length and to estimate growth in the model. Certainly, it is always important to analyse the age data with regard to these potential issues and to make sure that the data are appropriately stratified and scaled.

None of the issues relating to the problems of using conditional age-at-length data appear to have been considered in the cobia and Spanish mackerel assessments. The paucity of data is not an excuse for ignoring these issues – it does, in some ways, make it more important that they are considered.

Data weighting

There are various methods for obtaining relative weights (CVs and effective sample sizes) for the different data sets fitted in a stock assessment model. In both assessments, fairly arbitrary weights are used in the base models and iterative re-weighting methods (Francis 2011, SS3 re-weighting) are only considered in sensitivity runs. This is the wrong way round. The base runs should be using a formal weighting scheme and alternative schemes investigated in sensitivity runs. As it happens, it appears that the results are not particularly sensitive to the relative weights.

Effective sample sizes for composition data

This is partly covered under the data weighting heading (the method of Francis will give much lower sample sizes for composition data than SS3 re-weighting). However, in the cobia and mackerel assessments, the effective sample sizes that are used are based on the number of fish measured or aged (with a cap for sample size on length frequencies). This is not good practice. Best practice is to bootstrap the data to determine an effective sample size for each year based on how many fish were sampled in each trip and hence the within and between trip variability (and to use these sample sizes as initial values in iterative reweighting). Alternatively, if a rule-of-thumb is used, then the initial sample sizes should be based on the number of trips sampled rather than the total number of fish measured/aged. For example, if 100 fish were measured from 1 trip, the effective sample size should be closer to 1 than to 100 (e.g., Pennington et al. 2002). For age data the scaling down shouldn't be as extreme as for length data. For example, 100 fish aged from 10 trips could be worth 3-5 fish per trip, but almost certainly not 10 per trip.

That covers the joint problems.

Each of the ToRs are specifically considered below.

Cobia

1. *Evaluate the quality and applicability of data used in the assessment.*

Life history
The Life History Working Group covered the definition of stock boundaries and the estimation of fixed biological parameters. They considered appropriate data and made sensible recommendations with the exception of recommending 60% females at birth. They based this on the skewed sex ratios observed in the fisheries. However, the sex ratio in the population is hopelessly confounded with the fishing selectivities. It will make little difference, but the fishery dependent data considered do not give a reasonable basis to move from a 50-50 sex ratio at birth.

Catch history
The catch history was estimated for the commercial fishery starting in 1926 for three gear types (hand-line, long-line, and other). Recreational landings (which are much larger than the commercial landings) were calculated by mode and region (to some extent). Modes included charter-boat, headboat, private/rental boat, and shore based. Landings for Texas were calculated separately from the Gulf. Discard data for commercial and recreational fisheries were also compiled. The bycatch from the shrimp fishery, which was very substantial in some years, was also estimated (SEDAR28-DW6).

It is usually a difficult and tedious job to reconstruct full catch histories for stock assessment purposes and I think that a good job was done in this case. However, it would have been useful to provide the assessment team with an envelope of potential landings and discards so that they could have easily performed sensitivity runs with “low” and “high” levels of landings and discards.

Composition data
Available length and age data from the recreational and commercial fisheries were compiled by the Data Workshop (DW).

There was very little commercial length data and almost no commercial age data. The DW report says that the length data were "weighted by the landings in numbers by strata (state, year, gear)". This is not appropriate as many of the strata contained no samples. In order to get sensible length frequencies for the assessment there needed to have been an attempt to identify period of years which could be combined to provide adequate samples across a sensible post-stratification (e.g., combining some states). To determine an appropriate post-stratification requires an analysis of the variability of length frequencies across the various strata (e.g., it may be that some gear types could be combined). With so few samples the best that can probably be done is to construct a combined-year length frequency for each fishery.

The recreational sample sizes are also very low with many strata having zero or close to zero fish measured. Again it raises the issue of having to conduct a detailed analysis of the length data to determine how strata should be combined before scaling and production of annual or combined-year length frequencies. This is not discussed in the DW report at all so I must assume that no such analysis was done and that strata with low sample sizes (including zero) were just mechanically scaled.

Abundance indices

The Index Working Group (IWG) considered five potential abundance time series and recommended two of them for use in the assessment.

The SEAMAP data were not recommended because of the very low occurrence of cobia in the catch. A time series was developed from a delta-lognormal model. There is no mention in the DW report or the document they cite for details (SEDAR28-DW03) of why the indices were not constructed in the normal way for a trawl survey. Certainly, the original design was a random stratified trawl survey – so it makes no sense to use a delta-lognormal model which only measures density when abundance/biomass could have been measured. However, given the index was not used, my point is academic.

The Texas Parks and Wildlife Survey (TPWS) was analyzed using a delta-lognormal model where the data were restricted to an area that had relatively high cobia catches (SEDAR28-DW10). However, even for this area the proportion of positive trips was only 3.1% and the IWG did not recommend its use. The very low success rate does mean it is very unlikely to be tracking abundance.

A commercial vertical line index was constructed using the usual delta-lognormal model and no descriptive analysis at all (SEDAR28-DW16). The IWG did not recommend the time series because of the restrictive trip limit of two fish per person per day. The proportion of successful trips was also very low (2-4% each year). Certainly the derived indices could not be recommended. However, this dataset deserves more analysis. There may be a subset of trips which could provide some useful qualitative information on abundance from the proportion of positive trips.

The headboat and MRFSS datasets were analyzed to produce recreational CPUE indices (SEDAR28-DW28). Different filtering methods were considered and implemented but none were successful in identifying a subset of relevant cobia effort. Indices were calculated from just positive trips and also, using the delta-lognormal model, from all trips. Eventually the decision was made to base the index on all trips: "The working group also noted that there was little difference in the indices that were estimated for the entire dataset and the indices estimated for the subset of only positive trips. Therefore, it was reluctantly decided at the data workshop, that fishing effort for cobia and Spanish mackerel would be based on all trips".

I assume that the IWG felt that they had to recommend at least one time series for use as a relative abundance index in the stock assessment. However, the low level of successful trips for the headboat (7%) and MRFSS (<1%) datasets should have led to the same conclusion as for the TPWS. Additionally, there is the issue of the change in regulations in the period spanned by the time series and the different regulations in Florida state waters. These data may be able to provide useful abundance indices. However, an analysis based on an understanding of the various fisheries which occur over the region, will be needed to deliver defensible indices.

The two time series recommended by the IWG are not defensible in my opinion.

2. *Evaluate the quality and applicability of methods used to assess the stock.*

The stock assessment modeling was adequate but the assessment overall cannot recover from the poor data inputs. In the Data Workshop, there was inadequate attention to detail in regard to the composition data, and the recommended CPUE indices were not defensible as relative abundance indices.

Stock Synthesis 3

The Data Working Group recommended that the assessment be updated using ASPIC because of the paucity of composition data. This was a poor recommendation because the important fisheries for the stock have very different size/age based selectivities. It is not clear how the bycatch in the shrimp fishery could have been modeled satisfactorily in ASPIC or how a minimum legal size would have been implemented.

Perhaps an assessment could have been done in ASPIC, but then an equivalent assessment could also be done in SS3 – which can be run as an "age-based production model". The advantage of using SS3 is that there are numerous options for exploring the effect of fitting the available composition data and estimating or not estimating selectivity patterns and year class strengths.

Model structure

A simple and typical model structure was used. Population in age-structured equilibrium before the start of the fisheries. Year-round fisheries with constant selectivity patterns (with some time-blocking). Constant age-specific natural mortality over time. A single von Bertalanffy growth curve estimated in the model and a Beverton Holt stock-recruitment relationship. Year class strengths (recruitment deviations) were estimated from 1982-2010 (which is probably far too many given the paucity of composition data).

The shrimp fishery was modeled as a bycatch fishery with the catch driven by an effort time series and fitted to the median estimate of cobia bycatch from 1972-2011 using the "super-year" feature of SS3. Modeling the shrimp fishery in this way is a good approach.

Only a single commercial and a single recreational fishery were modeled despite the Data Working group providing landings histories for a number of fisheries. I assume the lumping of these data was because of the paucity of composition data but no explanation was provided in the Assessment Report. I have not considered whether it was justified or not – it would depend on whether the fisheries had similar selectivity patterns and whether their landings histories varied in a similar way over time.

Treatment of the data
The catch/landings histories were combined into single commercial and recreational fisheries which may or may not have been justified. The raw composition data, assembled by the Data Working group, seems to have been used in the assessment without any stratification or scaling (e.g., see Table 2.11 in the Assessment report – the number of fish measured is given in each year and then the number of fish in each 3cm bin is given; it looks like raw un-scaled data).

To get the most out of the limited composition data requires that it is very carefully post-stratified and scaled. The data are just there to help with estimation of growth, selectivities, and year class strength so it is unlikely to be fatal if they are not properly prepared; rather there is just a loss of information. Of course, if they are over-weighted relative to the abundance indices, then properly prepared or not they can severely distort an assessment.

The likelihood profile on virgin recruitment in the Assessment report (Figure 3.32) suggests that the age and length data are dominating the abundance indices in terms of a biomass signal (although it is a bit hard to tell – a "zoom in" would have been useful). The sample sizes, based on the number of fish measured or aged are too large. However, the abundance time series appear to be consistent with the biomass signal from the composition data so re-weighting of the data is unlikely to change the result.

Model runs
The base model used all of the available data and estimated steepness as well as numerous recruitment deviations. Given the paucity of composition data (and the fact it was not prepared properly) it is unlikely that there is good information on year class strength. The model will have no trouble coming up with estimates and will even provide good precision for those estimates because of the relatively high effective sample sizes assumed – but, in reality, the model is over-parameterized (and year class strengths are not well estimated).

Estimating steepness in these models is almost always the wrong thing to do. To get a good estimate requires excellent information on year class strengths over a wide range of relative spawning biomass. A glance at the available data tells us that steepness should not be estimated in this model.

A good range of sensitivity runs were performed, including low and high natural mortality and using one or other of the abundance time series. The only runs missing were those exploring the effects of different catch histories and discard rates. Certainly, the early catch history is very uncertain as are the discards from the shrimp fishery.

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

I cannot recommend any of the model runs for this assessment. The abundance indices are not defensible. The composition data were not properly prepared (and are over-weighted). The model was over-parameterized.

4. *Evaluate the methods used to estimate population benchmarks and management parameters. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.*

The methods used to estimate the SPR-based benchmarks are standard and done within SS3 which has been thoroughly tested. However, I cannot recommend any of the model runs and therefore do not provide any declarations of stock status.

5. *Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.*

   The base run and the low and high natural mortality runs were projected forward under three levels of fishing mortality ($F_{CURRENT}$, $FSPR_{30}$, and F_{OY}) using 1000 bootstrap replicates. The method is appropriate but I cannot recommend any of the runs.

6. *Evaluate the quality and applicability of methods used to characterize uncertainty in estimated parameters.*

   Uncertainty in the assessment was characterized by sensitivity runs and a parametric bootstrap on the base run. A good range of sensitivities were performed. The use of the bootstrap would not be my preferred choice but it is an acceptable approach. Calculation of Bayesian posteriors is generally preferable (even with uninformed priors). Also, uncertainty is badly under-estimated because of all the structural assumptions in the model (which is always the case) and the relatively large sample sizes used for the composition data (which does not have to be the case).

   - *Provide measures of uncertainty for estimated parameters*

     Confidence intervals from the bootstrap are provided in the Assessment report.

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

     The Assessment Report does not conclude that the assessment is highly uncertain and should be treated with extreme caution. This is my conclusion, mainly because of the lack of defensible abundance indices, but also because of the poor treatment of the composition data and the over-parameterization in the model.

   - *If there are significant changes to the base model, or to the choice of alternate states of nature, then provide a probability distribution function for the base model, or a combination of models that represent alternative states of nature, presented for review.*
     - *Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments*
     - *Provide justification for the weightings used in producing the combinations of models*

     Not applicable for this desktop review.

7. *If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.*

   Not applicable for this desktop review.

8. *Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.*

In general, the SEDAR process is a useful process for developing good quality stock assessments. However, the Data and Assessment Workshops in this case have not delivered good assessments.

Problems with the cobia assessment should have been identified at the Data Workshop – someone should have had the courage to say "we don't have a defensible abundance index" and they should have been listened to. The changes in fishing regulations and the variation between state and federal rules should have been noted by somebody.

The ToRs of the Data Workshop were each addressed. Of course, some were done better than others as I have already noted. The preparation of the composition data was very poor. The recommendation to scale the age data using the length frequencies was unfortunate.

ToR 5 for the Data Workshop requires them to recommend the assessment method. I don't think this is the role of a data workshop. They should get all the data together, in a form that provides options for the stock assessment (e.g., finer scale than that which might eventually be used in the stock assessment) but they shouldn't be telling the scientists who have to do the stock assessment modeling how to do it. Of course, ideally the person who has to do the modeling should be closely involved in all aspects of the Data Workshop.

The ToRs of the Assessment Workshop were each addressed. They used SS3 instead of ASPIC, which was a good choice. They didn't adequately document their reasons for some choices, such as using only a single commercial fishery and a single recreational fishery. They also appear to have used completely un-stratified and un-scaled composition data – certainly there is no explanation of how the data were scaled.

The review process normally involves a meeting where questions can be asked and answered and additional analyses used to explore issues. A desktop review, where the reviewers are not able to ask questions or discuss issues with the assessment scientists and each other, is not as good. Desktop reviewers only comment on the issues that they notice. In a meeting, issues that are noticed by each reviewer (and other meeting participants) come to the attention of all reviewers.

9. *Make any additional recommendations or prioritizations warranted.*
    - *Clearly denote research and monitoring needs that could improve the reliability of future assessments*

In the short-term, a new assessment is needed. There are no defensible abundance indices and it will hard to produce any quickly. Therefore, an assessment which looks at worst case scenarios should be considered. If the stock is in reasonable shape even at biomass levels that would only just allow the estimated catch to have been taken, then there is no rush to produce a full assessment.

Of course, a reliable assessment generally requires a defensible abundance time series. The development of such a series should be the top priority. Pursuit of such an index should also provide some answers on what other data need to be collected to provide defensible indices for cobia.

A workshop should be held to train people in the analysis and post-stratification of composition data.

Spanish Mackerel

10. *Evaluate the quality and applicability of data used in the assessment.*

Life history
The Life History Working Group covered the definition of stock boundaries and the estimation of fixed biological parameters. They considered appropriate data and made sensible recommendations with the exception of a strange recommendation on sex ratio: "Over all ages and gears, weighted percent females 66%". This was derived from their analysis of sex ratio data from fisheries. The Assessment Workshop took this as a recommendation for 50-50 at birth in 1886 (apparently): "Sex ratio at the start time of the population analysis (1886) was assumed to be 1:1 as recommended by the SEDAR 28 DW". It is strangely worded as 50-50 at birth in 1886 means 50-50 every year at birth.

Catch history
The catch history was estimated for the commercial fishery starting in 1880 for three gear types (gill nets, hand-line, and other). Recreational landings (which are much larger than the commercial landings) were calculated by mode and region (to some extent): MRFSS/MRIP estimates of landings from charter, private angler; Texas Parks and Wildlife (charter, private and headboat); and the for-hire headboat fishery. Discard data for commercial and recreational fisheries were also compiled. The bycatch from the shrimp fishery, which was very substantial in some years, was also estimated (SEDAR28-DW6).

It is usually a difficult and tedious job to reconstruct full catch histories for stock assessment purposes and I think that a good job was done in this case (no doubt building on the work done in previous assessments). However, it would have been useful to provide the assessment team with an envelope of potential landings and discards so that they could have easily performed sensitivity runs with "low" and "high" levels of landings and discards.

Composition data
Available length and age data from the recreational and commercial fisheries were compiled by the Data Workshop.

There were few commercial length and age data. The DW report says that the length data "were weighted by the trip landings in numbers and the landings in numbers by strata (state, year, gear).". This is not appropriate when many of the strata contained no samples. In order to get sensible length frequencies for the assessment there needed to have been an attempt to identify period of years which could be combined to provide adequate samples across a sensible post-stratification (e.g., combining some states). To determine an appropriate post-stratification requires an analysis of the variability of length frequencies across the various strata.

The recreational sample sizes are much higher but there are still a number of strata having zero or close to zero fish measured. Again it raises the issue of having to conduct a detailed analysis of the length data to determine how strata should be combined before scaling and production of annual or combined-year length frequencies. This is not discussed in the DW report at all so I must assume that no such analysis was done and that strata with low sample sizes (including zero) were just mechanically scaled. This is not a big issue for the MRFSS data, but for the headboat survey the sampling is very patchy and the data need to be carefully post-stratified.

Abundance indices

The Index Working Group (IWG) considered nine potential abundance time series and recommended three of them for use in the assessment.

The SEAMAP data were analyzed to produce an abundance time series for 0-1 year old Spanish mackerel (SEDAR28-DW03). The IWG recommended the time series for use because " it is a fisheries independent survey across a long time series (1987-2010), with very good spatial converge (TX/Mexico border to Mobile Bay)". Their statement is true but does not provide sufficient justification to include this time series in a stock assessment. In total, the two surveys each year caught between 32 and 487 fish. Typically, about 50-200 fish are caught each year. The proportion of positive stations was about 4% in summer and 8% in fall (SEDAR28-DW03). Basically, the survey doesn't catch much Spanish mackerel and the variability in the index is probably unrelated to the abundance of Spanish mackerel.

The three recreational surveys (Texas sport-boat angler survey, headboat, and MRFSS) all have very few successful trips. The IWG rejected the Texas and headboat surveys on this basis but recommended the use of the MRFSS time series although they didn't give any reasons other than: "This index was particularly favored because it presents a long time series." With less than 5% positive trips it is not reasonable to accept the unfiltered delta-lognormal time series as an abundance index.

Of the commercial data sets considered the IWG preferred the Florida State ticket data to the commercial logbook data for vertical lines and gillnets. I agree that the "run-around" gillnet method is likely to produce hyper-stable indices. Also, if Florida covers most of the fishery and has a longer time series then it is probably to be preferred to the shorter time series from the vertical line index (though, perhaps not in this case – see below).

The Florida trip-ticket data were used to construct cast net, hand-line/trolling, and gillnet indices split into time periods when trip limits were (assumed to be) not too restrictive. The IWG identified various problems with the "interpretation of data from trips using gill nets (e.g., deployment methods, mesh sizes, configuration of panels, and changes in state/federal waters restrictions) and cast nets (e.g., configuration, depth, bottom types)". I agree with their recommendation not to use these time series in stock assessment.

The IWG did recommend the Florida trip-ticket hand-line/trolling index (which shows an increasing trend over time) for use in stock assessment. This is a standardized index of catch-per-trip for trips that caught some Spanish mackerel (SEDAR28-AW01). The standardization approach is unusual as 8 of 11 explanatory variables are dummy variables which indicate whether a species-group was caught on the trip or not (this is slightly problematic as these are random variables and, strictly speaking, should not be used as explanatory variables). The remaining variables are year, month, and Florida sub-region. The documentation for this analysis does not mention using any measure of trip duration or "actual time fished" (which is a field on the Trip Ticket). They also do not make use of "number of crew" another field on the trip ticket (available since 2000). The response variable is given as "catch per trip" and not as “catch per trip per hour”. Perhaps this is just a documentation error? It is very hard to tell because there is no descriptive analysis to give a context to the standardization analysis. There is some discussion of outliers in the response variable: “those with landings greater than 1,223 pounds were excluded”. This tends to support “catch per trip”, but also it seems odd to exclude data on this basis – again the length of trip and the size of the vessel/number of crew, are important because longer trips and bigger vessels may catch more fish.

If “actual time fished” was not used in the standardization, and/or it is not properly reported on the form, then it is wrong to use this time series in stock assessment. The increasing trend could simply be the result of longer trips over time. It could also be the result of a change in the fleet with vessels that used to make short trips and/or not catch many fish, dropping out of the fishery over time. In a proper standardization these effects would be accounted for. It is also important when doing a standardization to first fully understand the data by doing a descriptive/exploratory analysis – it is very bad practice, as appears to have been done here, to simply “throw the data into the machine and turn the handle”. Not using “actual time fished” in the analysis is very hard to understand.

Unfortunately, I have found fatal faults with each of the three abundance times series used in the Spanish mackerel stock assessment.

11. *Evaluate the quality and applicability of methods used to assess the stock.*

The stock assessment modeling was adequate but the assessment overall cannot recover from the poor data inputs. In the Data Workshop, there was inadequate attention to detail in regard to the composition data, and the recommended CPUE indices were not defensible as relative abundance indices.

Stock Synthesis 3
The use of this package was appropriate given the available data.

Model structure
A simple and typical model structure was used. Population in age-structured equilibrium before the start of the fisheries. Year-round fisheries with constant selectivity patterns (with some time-blocking). Constant age-specific natural mortality over time. A single von Bertalanffy growth curve estimated in the model and a Beverton Holt stock-recruitment relationship. Year class strengths (recruitment deviations) were estimated from 1985-2010.

The shrimp fishery was modeled as a bycatch fishery with the catch driven by an effort time series and fitted to the median estimate of Spanish mackerel bycatch from 1972-2011 using the “super-year” feature of SS3. Modeling the shrimp fishery in this way is a good approach.

Two commercial fisheries were modeled but only a single recreational fishery was used despite the Data Working group providing landings histories for a number of fisheries. No explanation for this was provided in the Assessment Report. I have not considered whether it was justified or not – it would depend on whether the fisheries had similar selectivity patterns and whether their landings histories varied in a similar way over time.

Treatment of the data
The catch/landings histories were combined into two commercial fisheries and a single recreational fishery which may or may not have been justified. The raw length data, assembled by the Data Working group, seems to have been used in the assessment without

state in the stratification: "Length data were stratified by calendar year, fishery/survey (commercial gillnet fleet (COM_GN), commercial line gears (COM_RR), and recreational all fisheries combined (headboat, private angler, charter, shore = REC)". There should have been scaling from sample to trip and stratification needed to include state (unless there was an analysis showing that length frequencies were similar across states).

To get the most out of the limited composition data requires that it is very carefully post-stratified and scaled. The data are just there to help with estimation of growth, selectivities, and year class strength so it is unlikely to be fatal if they are not properly prepared; rather there is just a loss of information. Of course, if they are over-weighted relative to the abundance indices, then properly prepared or not they can severely distort an assessment.

The likelihood profile on virgin recruitment in the Assessment report (Figure 3.32) suggests that the age and length data are dominating the abundance indices in terms of a biomass signal (though it is a bit hard to tell – a "zoom in" would have been useful). The sample sizes, based on the number of fish measured or aged are too large. However, the abundance time series appear to be consistent with the biomass signal from the composition data so re-weighting of the data is unlikely to change the result.

Model runs

The base model (Run 3) used all of the available data and sensibly fixed steepness (0.8). Estimating steepness in these models is almost always the wrong thing to do. To get a good estimate requires excellent information on year class strengths over a wide range of relative spawning biomass.

A good range of sensitivity runs were performed, including low and high natural mortality and alternative values of steepness. The only runs missing were those exploring the effects of different catch histories and discard rates. Certainly, the early catch history is very uncertain as are the discards from the shrimp fishery.

12. *Recommend appropriate estimates of stock abundance, biomass, and exploitation.*

I cannot recommend any of the model runs for this assessment. The abundance indices are not defensible. The composition data were not properly prepared (and are over-weighted).

13. *Evaluate the methods used to estimate population benchmarks and management parameters. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.*

The methods used to estimate the SPR-based benchmarks are standard and done within SS3 which has been thoroughly tested. However, I cannot recommend any of the model runs and therefore do not provide any declarations of stock status.

14. *Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.*

The base run and a sensitivity run on steepness were projected forward deterministically under three levels of fishing mortality (FCURRENT, FSPR30, and F_{OY}). Stochastic projections using 1000 bootstrap replicates were also done for the base model. The method is adequate but I cannot recommend any of the runs.

15. *Evaluate the quality and applicability of methods used to characterize uncertainty in estimated parameters.*

    Uncertainty in the assessment was characterized by sensitivity runs and a parametric bootstrap on the base run. A good range of sensitivities were performed. The use of the bootstrap would not be my preferred choice but it is an acceptable approach. Calculation of Bayesian posteriors is generally preferable (even with uninformed priors). Also, uncertainty is badly under-estimated because of all the structural assumptions in the model (which is always the case) and the relatively large assumed sample sizes for the composition data (which does not have to be the case).

    - *Provide measures of uncertainty for estimated parameters*

      Confidence intervals from the bootstrap are provided in the Assessment report.

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

      The Assessment Report does not conclude that the assessment is highly uncertain and should be treated with extreme caution. This is my conclusion, mainly because of the lack of defensible abundance indices, but also because of the poor treatment of the composition data.

    - *If there are significant changes to the base model, or to the choice of alternate states of nature, then provide a probability distribution function for the base model, or a combination of models that represent alternative states of nature, presented for review.*
      - *Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments*
      - *Provide justification for the weightings used in producing the combinations of models*

      Not applicable for this desktop review.

16. *If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.*

    Not applicable for this desktop review.

17. *Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.*

    In general, the SEDAR process is a useful process for developing good quality stock assessments.

    The ToRs of the Data Workshop were each addressed. Of course, some were done better than others as I have already noted. The preparation of the composition data was poor. The recommendation to scale the age data using the length frequencies was very poor.

    ToR 5 for the Data Workshop requires them to recommend the assessment method. I don't think this is the role of a data workshop. They should get all the data together, in a form that provides options for the stock assessment (e.g., finer scale than that which might eventually be used in the stock assessment) but they shouldn't be telling the scientists who have to do the stock assessment modeling how to do it. Of course, ideally the person who has to do the modeling should be closely involved in all aspects of the Data Workshop.

The ToRs of the Assessment Workshop were each addressed. They didn't adequately document their reasons for some choices, such as using only a single recreational fishery. The stratification of the length data was very poor (state should have been included or a full justification given for ignoring it).

The review process normally involves a meeting where questions can be asked and answered and additional analyses used to explore issues. A desktop review, where the reviewers are not able to ask questions or discuss issues with the assessment scientists and each other, is not as good. Desktop reviewers only comment on the issues that they notice. In a meeting, issues that are noticed by each reviewer (and other meeting participants) come to the attention of all reviewers.

18. *Make any additional recommendations or prioritizations warranted.*
    - *Clearly denote research and monitoring needs that could improve the reliability of future assessments*

In the short-term, a new assessment is needed. There are data that may provide defensible abundance indices if analyzed properly (e.g., commercial logbook, vertical line data; Florida trip-ticket, hand-line/trolling data). It may also be possible to get something useful from the recreational data with appropriate filtering.

A workshop should be held to train people in the analysis and post-stratification of composition data.

Conclusions and Recommendations

The reviewed cobia and Spanish mackerel assessments are not suitable to be used to provide management advice because of the flawed data inputs used in the models.

My main conclusions are:

- Stock structure and fixed life history parameters were adequately considered.
- Landings history, discards, and discard mortalities were adequately determined and considered.
- Composition data were poorly treated at both the Data and Assessment Workshops. There was an absence of appropriate analysis and discussion with regard to post-stratification of the data to deal with inadequate sample sizes within some strata.
- The Index Working Group made very poor recommendations with regard to time series to use in the stock assessments as relative abundance indices:
    - For cobia, two recreational CPUE time series were recommended but these both had very low proportions of successful trips and spanned a period when fishing regulations had become more restrictive.
    - For Spanish mackerel: a SEAMAP survey was recommended as a recruitment time series, but it caught very few Spanish mackerel each year; a recreational time series was recommended but it had a very low proportion of successful trips; and a commercial index based on catch-per-trip was recommended but it had not been standardized for trip duration or time fished.
- None of the abundance indices used in the stock assessment runs were defensible.
- The model structure used, the choice of runs, and the methods of projection and capturing of uncertainty were adequate but could not overcome the flawed data inputs.
- None of the model runs should be used to determine biomass estimates or recommend stock status.

My main recommendations are:

- Top priority should be given to the construction of defensible abundance indices for both cobia and Spanish mackerel from the commercial and recreational data. I suggest the following approach:
    - Discussion with some of the participants in the fisheries to get some understanding of how, when, and where, they target cobia and Spanish mackerel.
    - A full descriptive/exploratory analysis of the data to understand the temporal and spatial variation in the catches and all of the available explanatory variables.
    - Identification of regional and seasonal fisheries for which fishing effort is likely to catch the species of interest (cobia or Spanish mackerel). This is likely to involve the identification of vessels in each year which fish at the times and places of interest and catch the species on some of their trips. It does not require that individual vessels be tracked across years (although that would be ideal).
    - An analysis to determine if fishing regulations have impacted on the ability of the data to track abundance (time series may have to be split to account for different fishing behaviour caused by regulation changes)
    - Production of standardized CPUE indices for each identified regional/seasonal fishery
    - Comparison of the trends across the different fisheries

    - Decide which if any of the CPUE indices are defensible as relative abundance indices (the length of the time series is not relevant to this decision).

- If defensible abundance indices can be constructed then assessments can be done as before except:
    - Composition data should be appropriately post-stratified and scaled; sample sizes should be based on the number of trips/landings sampled (not the number of fish measured or aged). This will require an analysis of the variability in length frequencies and proportion-at-age for given length across the various strata.
    - Recruitment deviates should only be estimated for cohorts which are well-represented in the composition data (e.g., appear at least three times in the age data).
    - Steepness should be fixed or estimated with an informed prior.

References

Chih, C. 2009. The effects of otolith sampling methods on the precision of growth curves. *North American Journal of Fisheries Management 29*: 1519-1528.

Cordue, P.L. 2008: Report on SEDAR 17, Stock Assessment Review, South Atlantic Vermilion Snapper and Spanish Mackerel, October 20 - 24, 2008, Savannah, Georgia. *For* CIE Independent System for Peer Review. 36 p.

Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. *Can. J. Fish. Aquat. Sci. 68*: 1124–1138.

Pennington, M.; Burmeister, L.; Hjellvik V. 2002. Assessing the precision of frequency distributions estimated from trawl-survey samples. *Fish. Bull. 100*: 74–80.

Stephens, A.; MacCall, A. 2004. A multispecies approach to sub-setting logbook data for the purposes of estimating CPUE. *Fisheries Research 70*: 299–310.

Venables, W.N.; Dichmont, C.M. 2004. GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research. *Fisheries Research 70*: 319–337.

Appendix 1: Bibliography of supplied material

The following data and assessment workshop reports were supplied for the desktop review.

SEDAR 28: Gulf of Mexico cobia, SECTION II: Data Workshop Report, May 2012. 239 p.
SEDAR 28: Gulf of Mexico Spanish mackerel, SECTION II: Data Workshop Report, May 2012. 268 p.
SEDAR 28: Gulf of Mexico cobia, SECTION III: Assessment Process Report, December 2012. 208 p.
SEDAR 28: Gulf of Mexico Spanish mackerel, SECTION III: Assessment Workshop Report, December 2012. 274 p.

The numerous workshop, background, and reference documents listed below were made available through the SEDAR website and were consulted as needed.

Document	Title	Authors
SEDAR28-DW01	Cobia preliminary data analyses – US Atlantic and GOM genetic population structure	Darden 2012
SEDAR28-DW02	South Carolina experimental stocking of cobia *Rachycentron canadum*	Denson 2012
SEDAR28-DW03	Spanish Mackerel and Cobia Abundance Indices from SEAMAP Groundfish Surveys in the Northern Gulf of Mexico	Pollack and Ingram, 2012
SEDAR28-DW04	Calculated discards of Spanish mackerel and cobia from commercial fishing vessels in the Gulf of Mexico and US South Atlantic	K. McCarthy
SEDAR28-DW05	Evaluation of cobia movement and distribution using tagging data from the Gulf of Mexico and South Atlantic coast of the United States	M. Perkinson and M. Denson 2012
SEDAR28-DW06	Methods for Estimating Shrimp Bycatch of Gulf of Mexico Spanish Mackerel and Cobia	B. Linton 2012
SEDAR28-DW07	Size Frequency Distribution of Spanish Mackerel from Dockside	N.Cummings, J. Isely

	Sampling of Recreational and Commercial Landings in the Gulf of Mexico 1981-2011	
SEDAR28-DW08	Size Frequency Distribution of Cobia from Dockside Sampling of Recreational and Commercial Landings in the Gulf of Mexico 1986-2011	J. Isely and N. Cummings
SEDAR28-DW09	Texas Parks and Wildlife Catch Per unit of Effort Abundance Information for Spanish mackerel	N. Cummings, J. Isely
SEDAR28-DW10	Texas Parks and Wildlife Catch Per unit of Effort Abundance Information for cobia	J. Isely, N. Cummings
SEDAR28-DW11	Size Frequency Distribution of Cobia and Spanish Mackerel from the Galveston, Texas, Reef Fish Observer Program 2006-2011	J Isely and N Cummings
SEDAR28-DW12	Estimated conversion factors for calibrating MRFSS charterboat landings and effort estimates for the South Atlantic and Gulf of Mexico in 1981-1985 with For Hire Survey estimates with application to Spanish mackerel and cobia landings	V. Matter, N Cummings, J Isely, K Brennen, and K Fitzpatrick
SEDAR28-DW13	Constituent based tagging of cobia in the Atlantic and Gulf of Mexico waters	E. Orbesen

SEDAR28-DW14	Recreational Survey Data for Spanish Mackerel and Cobia in the Atlantic and the Gulf of Mexico from the MRFSS and TPWD Surveys	V. Matter
SEDAR28-DW15	Commercial Vertical Line and Gillnet Vessel Standardized Catch Rates of Spanish Mackerel in the US Gulf of Mexico, 1998-2010	N. Baertlein, K. McCarthy
SEDAR28-DW16	Commercial Vertical Line Vessel Standardized Catch Rates of Cobia in the US Gulf of Mexico, 1993-2010	K. McCarthy
SEDAR28-DW17	Standardized Catch Rates of Spanish Mackerel from Commercial Handline, Trolling and Gillnet Fishing Vessels in the US South Atlantic, 1998-2010	K. McCarthy
SEDAR28-DW18	Standardized catch rates of cobia from commercial handline and trolling fishing vessels in the US South Atlantic, 1993-2010	K. McCarthy
SEDAR28-DW19	MRFSS Index for Atlantic Spanish mackerel and cobia	Drew et al.
SEDAR28-DW20	Preliminary standardized catch rates of Southeast US Atlantic cobia (*Rachycentron canadum*) from headboat data.	NMFS Beaufort
SEDAR28-DW21	Spanish mackerel preliminary data summary: SEAMAP-SA Coastal Survey	Boylan and Webster
SEDAR28-DW22	Recreational indices for cobia and Spanish mackerel in the Gulf of Mexico	Bryan and Saul
SEDAR28-DW23	A review of Gulf of Mexico and Atlantic Spanish mackerel (*Scomberomorus*	Palmer, DeVries, and Fioramonti

	maculatus) age data, 1987-2011, from the Panama City Laboratory, Southeast Fisheries Science Center, NOAA Fisheries Service	
SEDAR28-DW24	SCDNR Charterboat Logbook Program Data, 1993 - 2010	Errigo, Hiltz, and Byrd
SEDAR28-DW25	South Carolina Department of Natural Resources State Finfish Survey (SFS)	Hiltz and Byrd
SEDAR28-DW26	Cobia bycatch on the VIMS elasmobranch longline survey:1989-2011	Parsons et al.
SEDAR28-RW01	The Beaufort Assessment Model (BAM) with application to cobia: mathematical description, implementation details, and computer code	Craig
SEDAR28-RW02	Development and diagnostics of the Beaufort assessment model applied to Cobia	Craig
SEDAR28-RW03	The Beaufort Assessment Model (BAM) with application to Spanish mackerel: mathematical description, implementation details, and computer code	Andrews
SEDAR28-RW04	Development and diagnostics of the Beaufort assessment model applied to Spanish mackerel	Andrews
SEDAR28-RD01	List of documents and working papers for SEDAR 17 (South Atlantic Spanish mackerel) – all documents available on the SEDAR website	SEDAR 17
SEDAR28-RD02	2003 Report of the mackerel Stock Assessment Panel	GMFMC and SAFMC, 2003
SEDAR28-RD03	Assessment of cobia, Rachycentron canadum, in the waters of the U.S. Gulf of Mexico	Williams, 2001

SEDAR28-RD04	Biological-statistical census of the species entering fisheries in the Cape Canaveral area	Anderson and Gehringer, 1965
SEDAR28-RD05	A survey of offshore fishing in Florida	Moe 1963
SEDAR28-RD06	Age, growth, maturity, and spawning of Spanish mackerel, Scomberomorus maculates (Mitchill), from the Atlantic Coast of the southeastern United States	Schmidt et al. 1993
SEDAR28-RD07	Omnibus amendment to the Interstate Fishery Management Plans for Spanish mackerel, spot, and spotted seatrout	ASMFC 2011
SEDAR28-RD08	Life history of Cobia, Rachycentron canadum (Osteichthyes: Rachycentridae), in North Carolina waters	Smith 1995
SEDAR28-RD09	Population genetics of cobia Rachycentron canadum: Management implications along the Southeastern US coast	Darden et al, 2012
SEDAR28-RD10	Inshore spawning of cobia (Rachycentron canadum) in South Carolina	Lefebvre and Denson, 2012
SEDAR28-RD11	A review of age, growth, and reproduction of cobia Rachycentron canadum, from US water of the Gulf of Mexico and Atlantic ocean	Franks and Brown-Peterson, 2002
SEDAR28-RD12	An assessment of cobia in Southeast US waters	Thompson 1995
SEDAR28-RD13	Reproductive biology of cobia, Rachycentron canadum, from coastal waters of the southern United States	Brown-Peterson et al. 2001
SEDAR28-RD14	Larval development, distribution, and ecology of cobia Rachycentron canadum (Family: Rachycentridae) in the northern Gulf of Mexico	Ditty and Shaw 1992

SEDAR28-RD15	Age and growth of cobia, Rachycentron canadum, from the northeastern Gulf of Mexico	Franks et al 1999
SEDAR28-RD16	Age and growth of Spanish mackerel, Scomberomorus maculates, in the Chesapeake Bay region	Gaichas, 1997
SEDAR28-RD17	Status of the South Carolina fisheries for cobia	Hammond, 2001
SEDAR28-RD18	Age, growth and fecundity of the cobia, Rachycentron canadum, from Chesapeake Bay and adjacent Mid-Atlantic waters	Richards 1967
SEDAR28-RD19	Cobia (Rachycentron canadum) tagging within Cheasapeake Bay and updating of growth equations	Richards 1977
SEDAR28-RD20	Synopsis of biological data on the cobia Rachycentron canadum (Pisces: Rachycentridae)	Shaffer and Nakamura 1989
SEDAR28-RD21	South Carolina marine game fish tagging program 1978-2009	Wiggers, 2010
SEDAR28-RD22	Cobia (Rachycentron canadum), amberjack (Seriola dumerili), and dolphin (Coryphaena hipurus) migration and life history study off the southwest coast of Florida	MARFIN 1992
SEDAR28-RD23	Sport fish tag and release in Mississippi coastal water and the adjacent Gulf of Mexico	Hendon and Franks 2010
SEDAR28-RD24	VMRC Cobia otolith preparation protocol	VMRC
SEDAR28-RD25	VMRC Cobia otolith ageing protocol	VMRC
SEDAR28-RD26	Age, growth, and reproductive biology of greater amberjack and cobia from Louisiana waters	Thompson et al. 1991

SEDAR28-RD27	Gonadal maturation in the cobia, Rachycentron canadum, from the northcentral Gulf of Mexico	Lotz et al. 1996
SEDAR28-RD28	Cobia (Rachycentron canadum) stock assessment study in the Gulf of Mexico and in the South Atlantic	Burns et al. 1998
SEDAR28-RD29	Total mortality estimates for Spanish mackerel captured in the Gulf of Mexico commercial and recreational fisheries 1983 to 2011	Bryan 2012

SEDAR28-AW01	Florida Trip Tickets	S. Brown
SEDAR28-AW02	SEDAR 28 Spanish mackerel bycatch estimates	NMFS Beaufort

Appendix 2: Statement of Work for Patrick Cordue

Amended Statement of Work

External Independent Peer Review by the Center for Independent Experts

SEDAR 28: Gulf of Mexico Cobia and Spanish Mackerel Assessment Desk 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 Representative (COR), 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 28 will be a compilation of data, an assessment of the stocks, and an assessment review conducted for Gulf of Mexico Spanish mackerel and cobia. The CIE peer review is ultimately responsible for ensuring that the best possible assessment has been provided through the SEDAR process. The stocks assessed through SEDAR 28 are within the jurisdiction of the Gulf of Mexico Fisheries Management Councils and states in the Gulf of Mexico region. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements for CIE Reviewers: Three CIE reviewers shall have the necessary qualifications to complete an impartial and independent peer review in accordance with the statement of work (SoW) tasks and terms of reference (ToRs) specified herein. The CIE reviewers shall have expertise in stock assessment, statistics, fisheries science, and marine biology sufficient to complete the tasks of the peer-review described herein. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall participate and conduct an independent peer review as a desk review, therefore travel will not be required.

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 contact information to the COR, 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 assessment and other pertinent background documents for the peer review. Any changes to the SoW or ToRs must be made through the COR 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.

Desk Review: 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 shall not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COR and CIE Lead Coordinator. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review 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.

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) Conduct an impartial and independent peer review in accordance with the tasks and ToRs specified herein, and each ToRs must be addressed (**Annex 2**).
3) No later than January 25, 2013, 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 Dr. 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.

21 December 2012	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
2 January 2013	NMFS Project Contact sends the CIE Reviewers the assessment report and background documents
9-24 January 2013	Each reviewer conducts an independent peer review as a desk review
25 January 2013	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
8 February 2013	CIE submits CIE independent peer review reports to the COR
15 February 2013	The COR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on changes. The COR 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 COR 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 COR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) The CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) The CIE report shall address each ToR as specified in **Annex 2**,
(3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

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

Support Personnel:

William Michaels, Program Manager, COR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-427-8155

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:

Ryan Rindone, SEDAR Coordinator
2203 N. Lois Avenue, Suite 1100
Tampa, FL 33607
Ryan.Rindone@gulfcouncil.org Phone: 813-348-1630

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.

   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

Annex 2a – Terms of Reference for

SEDAR 28: Gulf of Mexico Cobia Assessment Desk Review

1. Evaluate the quality and applicability of data used in the assessment.
2. Evaluate the quality and applicability 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. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.
5. Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.
6. Evaluate the quality and applicability 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
    - If there are significant changes to the base model, or to the choice of alternate states of nature, then provide a probability distribution function for the base model, or a combination of models that represent alternative states of nature, presented for review.
        - Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments
        - Provide justification for the weightings used in producing the combinations of models
7. If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.
8. Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.
9. Make any additional recommendations or prioritizations warranted.
    - Clearly denote research and monitoring needs that could improve the reliability of future assessments

Table 1. Required MSRA Evaluations for cobia assessment:

Criteria	Definition* (2001)	Current Value* (2001)
Mortality Rate Criteria		
F_{MSY}	F_{MSY}	0.34
MFMT	F_{MSY}	0.34
F_{OY}	75% of F_{MSY}	0.26
$F_{CURRENT}$	F_{2000}	0.30
$F_{CURRENT}$/ F_{MSY}	Percentage of $F_{Current}/F_{MSY}$ > MFMT	0.40
Base M		0.30
Biomass Criteria		
SSB_{MSY}	Equilibrium SSB_{MSY} @ F_{MSY}	3.02 mp
MSST	(1-M)*SSB_{MSY}: M=0.30	2.11 mp
$SSB_{CURRENT}$	SSB_{2000}	
$SSB_{CURRENT}$/ SSB_{MSY}	Percentage of $SSB_{Current}/SSB_{MSY}$ < MSST	0.30
Equilibrium MSY	Equilibrium Yield @ F_{MSY}	1.50 mp
Equilibrium OY	Equilibrium Yield @ F_{OY}	1.45 mp
OFL	Annual Yield @ MFMT	
	2013	
	2014	
	2015	
	2016	
	2017	
	2018	
Annual OY**	Annual Yield @ F_{OY}	
	2013	
	2014	
	2015	
	2016	
	2017	
	2018	

*Definitions and values are subject to change as per guidance from this assessment.

**Based upon current definitions of OY, where OY = 75% of F_{MSY}

Table 2. Projection Scenario Details for cobia assessment

2.1 Initial Assumptions:

OPTION	Value
2012 base TAC	TBD
2012 Recruits	TBD by Panel
2012 Selectivity	TBD by Panel
Projection Period	6 yrs (2013-2018)
1st year of change F, Yield	2013

2.2 Scenarios to Evaluate (preliminary, to be modified as appropriate)

1. Landings fixed at 2013 target
2. F_{OY}= 65%, 75%, 85% F_{MSY} (project when OY will be achieved)
3. F_{MSY}
4. $F_{REBUILD}$ (if necessary)
5. F=0 (if necessary)

2.3 Output values

1. Landings
2. Discards (including dead discards)
3. Exploitation
4. F/F_{MSY}
5. B/B_{MSY}

Annex 2b – Terms of Reference for

SEDAR 28: Gulf of Mexico Spanish Mackerel Assessment Desk Review

10. Evaluate the quality and applicability of data used in the assessment.
11. Evaluate the quality and applicability of methods used to assess the stock.
12. Recommend appropriate estimates of stock abundance, biomass, and exploitation.
13. Evaluate the methods used to estimate population benchmarks and management parameters. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.
14. Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.
15. Evaluate the quality and applicability 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
    - If there are significant changes to the base model, or to the choice of alternate states of nature, then provide a probability distribution function for the base model, or a combination of models that represent alternate states of nature, presented for review.
        - Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments
        - Provide justification for the weightings used in producing the combinations of models
16. If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.
17. Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.
18. Make any additional recommendations or prioritizations warranted.
    - Clearly denote research and monitoring needs that could improve the reliability of future assessments

Table 1. Required MSRA Evaluations for Spanish mackerel assessment:

Note: te = trillion eggs

Criteria	Definition* (as of 2002/2003)	Current Value* (2002/03)
Mortality Rate Criteria		
F_{MSY}	$F_{30\%SPR}$	
MFMT	$F_{30\%SPR}$	
F_{OY}	75% of $F_{30\%SPR}$	0.40
$F_{CURRENT}$	$F_{2002/03}$	
$F_{CURRENT}$/MFMT		0.53
Base M		0.30
Biomass Criteria		
SSB_{MSY}	Equilibrium SSB_{MSY} @ $F_{30\%SPR}$	19.10 te
MSST	(1-M)*SSB_{MSY}: M=0.30	13.40 te
$SSB_{CURRENT}$	SSB_{2003}	17.96 te
$SSB_{CURRENT}$/ MSST		1.34
Equilibrium MSY	Equilibrium Yield @ $F_{30\%SPR}$	8.7 mp
Equilibrium OY	Equil. Yield @ 75% of $F_{30\%SPR}$	8.3 mp
OFL	Annual Yield @ MFMT	
	2013	
	2014	
	2015	
	2016	
	2017	
	2018	
Annual OY**	Annual Yield @ F_{OY}	
	2013	
	2014	
	2015	
	2016	
	2017	
	2018	

*Definitions and values are subject to change as per guidance from this assessment.

**Based upon current definitions of OY, where OY = 75% of F_{MSY}

Table 2. Projection Scenario Details for Spanish mackerel assessment

2.1 Initial Assumptions:

OPTION	Value
2012 base TAC	TBD
2012 Recruits	TBD by Panel
2012 Selectivity	TBD by Panel
Projection Period	6 yrs (2013-2018)
1^{st} year of change F, Yield	2013

2.2 Scenarios to Evaluate (preliminary, to be modified as appropriate)

1. Landings fixed at 2013 target
2. F_{OY}= 65%, 75%, 85% F_{MSY} (project when OY will be achieved)
3. F_{MSY}
4. $F_{REBUILD}$ (if necessary)
5. F=0 (if necessary)

2.3 Output values

1. Landings
2. Discards (including dead discards)
3. Exploitation
4. F/F_{MSY}
5. B/B_{MSY}

Report on the SEDAR 28 Desk Review of the Stock Assessments for Gulf of Mexico Cobia and Spanish Mackerel

N. G. Hall

Unit 2
2 Wexford Street
Subiaco
Western Australia 6008
Australia
Phone: +61 8 9401 6891
Email: nghall@iinet.net.au

February 2013

Contents

1. Executive Summary

Between 9 and 24 January 2013, a Center for Independent Experts (CIE) desk review of the SEDAR 28 Gulf of Mexico cobia (*Rachycentron canadum*) and Spanish mackerel (*Scomberomorus maculates*) stock assessments was undertaken. The key findings of that review are summarised below.

Prior to the development of assessment models by the Assessment Workshops, the Data Workshops had collated the biological data for the Gulf of Mexico stocks of cobia and Spanish mackerel and constructed time series of reliable data for the landings made by the commercial and recreational fisheries. Despite some deficiencies of the data collection programs, the Workshops had developed time series of discards from these fisheries and of the bycatch of the two species from the shrimp fishery. Although imprecise, these time series, together with the time series of landings data, had been considered appropriate for use in the assessments. Length composition data sufficient to characterize the landings data, and, in the case of the Spanish mackerel stock, one of the survey indices, had been collated, together with those age-at-length data that were available. The Data Workshop for cobia had also recommended two fishery-dependent survey indices, while that for Spanish mackerel had recommended one fishery-independent index of abundance and two fishery-dependent indices. Each of the survey indices had been standardized using an appropriate statistical approach.

Although both maturity at age and the various time series of discard data for both species were imprecise, and there was a lack of length and age-at-length composition data for those fish that had been discarded from the commercial and recreational fisheries, the data that the Data Workshops had collated for the Gulf of Mexico stocks of both cobia and Spanish mackerel represented the best data that were available and were considered adequate for use in stock assessment. It should be noted, however, that the imprecision of the input data and limited age composition data are reflected in uncertainty in the results of each assessment. In the case of cobia, the lack of a fishery-independent index of abundance is also likely to have influenced the results that were obtained from the assessment.

Assessments for both cobia and Spanish mackerel had been undertaken by the Assessment Workshops using Stock Synthesis 3, a versatile and well-tested program that has been employed in numerous stock assessments both in the U.S. and elsewhere. The methods employed by this program are of high quality and the software provides tools that facilitate exploration of uncertainty, calculation of benchmarks, projection of yields with specified fishing rates to assess future stock status, and, through bootstrapping, either within Stock Synthesis (in the case of cobia) or using auxiliary software (in the case of Spanish mackerel), generation of probability distributions of parameters, benchmarks, and other variables. The ease with which alternative values of parameters can be set up within Stock Synthesis had facilitated (1) the exploration by the Assessment Workshops of the sensitivity of the results produced by the cobia and Spanish mackerel models to a number of alternative assumptions regarding values of natural mortality, steepness, and discard mortality, (2) the conducting of retrospective analyses, and (3) investigation of alternative data weighting options.

For both cobia and Spanish mackerel, estimates of the steepness of the stock-recruitment relationship had been found to be imprecise. The key uncertainty reflected in the choice by the Assessment Workshop for Gulf of Mexico cobia of a set of models to

represent alternative states of nature was the value of steepness. For Spanish mackerel, the Assessment Workshop chose to explore the effects of a range of values for the base level of natural mortality *M* when proposing alternative states of nature. Sensitivity analysis had also indicated that the results of the assessment for cobia were sensitive to this parameter.

The base model for the Gulf of Mexico stock of cobia assumed a base level of natural mortality of 0.38 y^{-1}, which, when fitted, resulted in an estimated steepness of 0.925. Based on the sensitivity analyses and explorations of uncertainty that had been carried out by the Assessment Workshop, this model and two alternative models were accepted as suitable for use as alternative states of nature when assessing the condition of the cobia stock. The alternative models assumed base levels of natural mortality of 0.26 and 0.5 y^{-1}, and, when fitted, resulted in steepness estimates of 0.96 and 0.92, respectively. On fitting the base model for the Gulf of Mexico stock of cobia, it was estimated that SSB_{2011}/MSST=1.73 and that F_{current}/ MFMT = 0.63, where the benchmarks MSST and MFMT had been calculated as MFMT = $F_{30\%SPR}$ and MSST = $(1 - M)$ $SSB_{30\%SPR}$. Based on this result and the examination of the results of the various sensitivity runs for Gulf of Mexico cobia, it is highly likely that the stock of cobia is not overfished and is not experiencing overfishing.

Exploration of parameter estimates, sensitivity runs, likelihood profiles, and results from bootstrapping led the Assessment Workshop for the Gulf of Mexico stock of Spanish mackerel to accept an alternative to the initial model as the new base model for this species. While this new model had an identical structure to that of the original base model, the value of steepness was fixed at 0.8, rather than estimated. An alternative model with similar structure to that of the new base model, but with steepness fixed at 0.9, was chosen by the Assessment Workshop to represent an alternative state of nature. Estimates obtained from the fitted base model indicated that SSB_{2011}/MSST=3.06 and that F_{current}/ MFMT = 0.38, where the benchmarks MSST and MFMT had been calculated as MFMT = $F_{30\%SPR}$ and MSST = $(1 - M)$ $SSB_{30\%SPR}$. Based on this result and examination of the results of the various sensitivity runs, it is highly likely that the Gulf of Mexico stock of Spanish mackerel is not overfished and is not experiencing overfishing.

The assessments produced by the Assessment Workshops for the Gulf of Mexico stocks of cobia and Spanish mackerel are based on the best data that are available, and the models that have been developed in Stock Synthesis are appropriate given the input data that are available for each stock. The results of these assessments provide the best scientific advice regarding the status of these two stocks that is currently available. While the limitations of the data and the uncertainty reflected in the sensitivity analyses and in the values calculated by the assessment models should be recognized when considering future management options, the explorations described in the Assessment Workshop Reports suggest that the conclusions regarding current stock status and levels of fishing mortality are likely to be robust despite the uncertainty associated with the assessments. Future stock assessments would benefit from improvement in the programs used (1) to collect discard data from the commercial and recreational fisheries and bycatch data from the shrimp fishery, and (2) to collect length and age-at-length data from landings and discards from both the commercial and recreational fisheries and from the bycatch of cobia and Spanish mackerel by the shrimp fishery.

The individuals involved in collating the input data and in developing the stock assessments are commended for their efforts.

2. Background

2.1. Overview

Between 9 and 24 January, 2013, a Center for Independent Experts (CIE) desk review was undertaken of the SEDAR 28 Gulf of Mexico cobia and Spanish mackerel stock assessments.

The Statement of Work provided to Dr Norm Hall by the CIE is attached as Appendix 2. This CIE report, which is prepared in accordance with the Statement of Work, describes his evaluation of the assessments and the review process.

Prior to the Review, stock assessment documents and other background documentation were made available to CIE Reviewers. A list of these documents is presented in Appendix 1. Note that, in the text of this review report, the "Gulf of Mexico – Cobia – Assessment Process Report" is referred to as the "Workshop Assessment Report" for the Gulf of Mexico stock of cobia.

2.2. Terms of Reference

The terms of reference for the desk review of the stock assessments of the Gulf of Mexico stocks of cobia and Spanish mackerel are presented in the Statement of Work (Appendix 2).

3. Description of Reviewer's role in review activities

Prior to undertaking the desk review, the Reviewer familiarised himself with the background documentation and the assessment reports for the two species that were the subject of the review (Appendix 1). Subsequently, he examined the Data Workshop and Assessment Workshop Reports for each species in greater detail, focussing on the preparation of this document, *i.e.*, the CIE report describing his evaluation of the two stock assessments and the SEDAR process.

4. Summary of findings relevant to the SEDAR 28 stock assessments for Gulf of Mexico cobia and Spanish mackerel

Because of the similarity of the models and many aspects of the data for the Gulf of Mexico stocks of cobia and Spanish mackerel, common issues in both assessments were often identified. There is thus some duplication of the text used when discussing those issues under the Terms of Reference for the separate stocks.

*4.1 Gulf of Mexico Cobia (*Rachycentron canadum*).*

ToR 1. Evaluate the quality and applicability of data used in the assessment.

Conclusions

The data that the Data Workshop has compiled for the Gulf of Mexico stock of cobia are the best that are available. Although limited, and imprecise in some aspects, the data are of a quality that allows a broad assessment of the likely condition of the stock.

Strengths

- The collation of life history data for the Gulf of Mexico stock of cobia.
- The collation of commercial landings data to produce time series of landings by handline, longline, and other gears from 1927, and, particularly, more precise data from 1950.
- The collation of a time series of estimates of bycatch of cobia by the shrimp fishery from 1972, using a Bayesian model to estimate catch per unit of effort.
- The collation of recreational fisheries data from different sources to produce sound time series of landings by fishing mode from 1955, and, particularly, more precise data from 1981.
- The collation of data to produce time series of discards from the commercial gears and recreational fishing modes.
- The collation of length composition data to characterize the landings by the commercial and recreational fisheries.
- The collation of two fishery-dependent indices of abundance, and the use of appropriate statistical analyses to standardize those indices of abundance.

Weaknesses

- Lack of definition of the southern boundary of the Gulf of Mexico stock of cobia.
- Paucity of data on the relationship of the proportion mature with age.
- The unreliable nature of the discard data due to low reporting, low intercept rates, and inadequate data collection programs.
- Inadequate sampling of length and age composition data from commercial landings and from bycatch of cobia from the shrimp fishery.

- Lack of length and age composition sampling from commercial and recreational discards.

Specific comments

Stock structure

The decision that, during the spawning season, mature individuals of cobia in the Gulf of Mexico are genetically distinct from those on the Atlantic coast north of Florida appears sound given the genetic and tagging data that are available. While the number of cobia in the sample collected in waters off Texas for the genetic study appears adequate, samples from the north of the Gulf of Mexico and from waters off the west coast of Florida are small. Further research to collect additional data from within the Gulf and to confirm the preliminary genetic findings would be valuable.

Despite the overall conclusion that the Gulf of Mexico stock is distinct from the South Atlantic stock of cobia, the genetic and tagging data indicate that there is some gene flow and a small amount of movement between the stock in the Gulf and those stocks in the stock complex off the South Atlantic coast, the latter complex being considered as the South Atlantic "stock" of cobia. There is also an inconsistency between the findings reported in SEDAR28-DW01 and those reported in SEDAR28-RD09, which needs to be reconciled. The former report advises that the collections from offshore in the Gulf of Mexico were genetically distinct from those offshore in the South Atlantic region, while the latter reports that "Based on our U.S. collections of *R. canadum* encountered along the SA and GOM coasts, tests of both genotypic distributions and pairwise hierarchical RST statistics suggest the offshore groups are genetically homogenous, even between the SA and GOM" and that "information gathered from the offshore collections … shows high levels of movement between the SA and GOM".

From the Data Workshop Report, it appears that the majority of tag recoveries have been made in locations that are consistent with the location of release of the tagged fish and the results of genetic studies of fish collected during the spawning season. Although not stated in this Report, the temporal distribution of recaptures of tagged fish presumably reflects the temporal distribution of catches in both spawning and non-spawning periods. The tag recovery data thus suggest that, despite the migrations that cobia undertake, regardless of the time of year and with the exception of fish caught in the waters off Brevard County, catches of fish may be assigned reliably to one or other of the two stocks on the basis of the area in which they are caught. Genetic studies should be undertaken to confirm this hypothesis, however.

As concluded in the Data Workshop Report, the genetic and tagging data indicate that Gulf of Mexico and South Atlantic stocks of cobia overlap in the waters to the east of Florida, and there is thus no distinct boundary that separates the stocks. For assessment and management, and for allocation of catches to one or other of the two stocks, the boundary between Florida and Georgia was selected (for convenience and because it was consistent with genetic, tagging and life history data) as the line separating the two stocks. Consideration should be given to whether catches within the area of overlap are of sufficient magnitude that assessment results could be sensitive to this decision, *i.e.*, whether an assessment based on an alternative line of separation at, say, the southern edge of the zone of overlap of the two stocks would be likely to yield results that differ greatly from those reported for the current assessment.

Unfortunately, maps of the distribution of the species and stocks of cobia, which were requested in the terms of reference for the Data Workshop, were not prepared. FishBase (Froese and Pauly, 2012) advises, however, that cobia has a worldwide distribution, which extends south of U.S. waters into waters off South America. The genetic study provides no information to suggest that the Gulf of Mexico stock does not extend into waters off Mexico, where it may also experience the effects of fishing. Further genetic research to determine the southern extent of the Gulf of Mexico stock of cobia appears necessary.

Biological data

The Life History Working Group's recommendation to base its estimate of the average value of the instantaneous rate of natural mortality *M* for fully-selected fish (ages 3-11) on the value determined from the Hoenig (1983) equation for fish using a maximum age of 11 years, *i.e.*, 0.38 y^{-1}, is endorsed. The range of estimates of *M* ultimately used to explore the sensitivity of the assessment model to imprecision in the estimate of natural mortality, *i.e.*, 0.26 to 0.5 y^{-1}, was broader than that initially proposed by the Life History Working Group (LHWG), *i.e.*, 0.26 to 0.42 y^{-1}. While the LHWG also recommended that a range of values of *M* based on a CV of 0.54 (MacCall, 2011), or other CVs, should also be explored, such exploration does not appear to have been undertaken by the Assessment Workshop. The basis for the use of 0.5 y^{-1} as a high value of *M* is not explained in the Assessment Workshop Report, but it is noted that the difference between this high value and the base level of 0.38 y^{-1} is equal to the difference between that latter value and the low value of 0.26 y^{-1}. Research is needed to determine methods by which an appropriate range of feasible values of *M* for a species might be selected for use in stock assessment as alternate plausible states of nature.

For Gulf of Mexico cobia, estimates of *M* from the Lorenzen equation were scaled such that the average value of *M* over the fully-selected ages 3 to 11 years was equal to the estimate from Hoenig's (1983) equation for fish, *i.e.*, 0.38 y^{-1}. It is unclear, however, whether the same approach as used for Run 1 was applied in sensitivity runs 2 and 3 when, as advised in the Assessment Workshop Report, the Lorenzen-based age dependent mortalities were scaled to achieve the same cumulative survivals over all ages as that expected for constant mortalities equal to the low and high values of *M*, respectively. It is likely that the cumulative survival was calculated over only ages 3-11, rather than all ages, to ensure consistency with the approach used in Run 1 when average *M* was set to 0.38 y^{-1}.

Use of the Lorenzen (1996) equation to derive age-dependent estimates of natural mortality *M* is not endorsed. In his report to the CIE on the stock assessments conducted for yellowtail flounder and Atlantic herring at Woods Hole in 2012, Francis (2012) advised that prediction of *M*, and, through body weight, its variation with age for an individual species, using Lorenzen's (1996) equation was likely to be highly imprecise, as was evident in the wide scatter about the regression line in Lorenzen's Figure 1. Francis observed that, for about one-third of Lorenzen's data points, predicted and observed *M*s appeared to differ by a factor of more than 2. Furthermore, in the case of both herring and yellowtail, the values of *M* estimated by Lorenzen's (1996) equation differed markedly from the values estimated using Hoenig's (1983) equation and had to be scaled substantially for use in the yellowtail flounder and Atlantic herring assessments. If it is assumed that the length measure used for Gulf of Mexico cobia in the growth equation, the parameters of which are presented in Table 2.7.1 of the Data Workshop Report, is fork length rather than total length (not advised in the text or table but inferred from Fig. 2.7.2), the value of *M* at age 3 is estimated by the Lorenzen (1996) equation to be 0.21 y^{-1}. This suggests that the estimates for the Gulf of Mexico stock of cobia calculated using Lorenzen's (1996) method were scaled up by a factor of at least 1.8 to produce the estimates of age-dependent natural mortality used in the assessment. Francis (2012) raised the valid point that, if the estimates produced for a species by Lorenzen's (1996) equation provide such unreliable estimates that the mean *M* differs from the estimate calculated using Hoenig's (1983) equation by a factor that differs markedly from 1, can it be considered sufficiently reliable to estimate how *M* varies with age within these species?

There has been no test to assess whether the introduction of the additional complexity associated with age-dependent natural mortality was justified by the resultant improvement in fit that was obtained for the Gulf of Mexico cobia model. It is recommended that a model employing a constant value of *M* is fitted to the cobia data. If this model fits just as well as the model that employs an age-dependent *M*, then the simpler model should be used. If the age-dependent model produces a significantly better fit, it would probably be better to estimate age-dependent *M* within the assessment model rather than assuming that it is of the form predicted by the Lorenzen (1996) equation.

The Data Workshop's decision, that cobia are hardy and unlikely to suffer barotrauma-associated post-release mortality, is subjective. Further research on discard mortality would be useful.

The Data Workshop correctly identified that, because of bias introduced into biological samples by the 33 inch minimum legal size, an allowance would need to be made when fitting von Bertalanffy growth curves to length-at-age data. By fitting the growth curves in Stock Synthesis, the influence of the selection curves on the observed length-at-age data is automatically taken into account and uncertainty associated with fitting the growth curves is carried through to the estimates of parameters and benchmarks that are produced by Stock Synthesis.

Because of the paucity of the youngest ages of fish in samples, the advice relating to maturity at age, which was reported in the Data Workshop Report, was subjective. Research based on fishery-independent samples is needed to provide more reliable estimates of the parameters of the maturity-length relationship and the proportion mature at age.

Although the Data Workshop noted that cobia exhibit sexually dimorphic growth, the Stock Synthesis model used in the assessment employed only the growth curve for the pooled sexes. In future refinement of the assessment model, consideration should be given to including sexually dimorphic growth, noting that the benefit of this might only be realised if appropriate sex composition data for landings and discards become available for input, and length and age-at-length compositions are sexually disaggregated.

Commercial landings

The decision by the Data Workshop to extend the historical time series of commercial landings of Gulf of Mexico cobia as far as possible into the past is endorsed, as catches from that earlier time period are likely to have influenced current stock status. It was noted that the Data Workshop reported that "Landings prior to 1950 are considered highly uncertain" and that the precision of landings improved following the introduction of the trip ticket system in each state. The tables that are presented provide no estimates of the precision likely to be associated with the annual landings data, nor is any information provided as to whether the commercial landings for cobia, which were reported by the Data Workshop, were likely to be biased, and, if so, the magnitude and direction of such bias.

Without an alternative time series, such as fishing effort, to provide information on fishing mortality, Stock Synthesis assumes that the catches are known sufficiently well to estimate the fishing mortalities required to take those catches (Methot and Wetzel, 2012), and thus estimated catches match the values that were input. In the current assessment, there has been no evaluation of the implications of the greater imprecision of the commercial landings data prior to 1950. Such evaluation may have required a sensitivity run with an alternative time series of commercial landings encompassing the imprecision of the landings data.

The Data Workshop has reported that, because few trips with cobia discards were observed by the Reeffish Observer Program and the NMFS logbook does not provide coverage of the entire fishery, discards of cobia by the commercial fishery have greater uncertainty than commercial landings and are likely to underestimate the true quantities of discarded fish. No estimate is provided of the likely magnitude of such underestimation.

The Working group advised that discards reported as "kept, not sold" should be added to the landings, and not included in the discards. This recommendation does not appear to have been accepted by the Assessment Workshop as Table 3.6 of the Data Workshop Report includes these fish within the discards, and the same values are carried over and used in the assessment (Table 2.5 and Appendix A, Assessment Workshop Report). The value for 2011 in Table 2.5 differs from that reported in Appendix A in the Assessment Workshop Report.

The estimates of the annual bycatch of cobia in the Gulf of Mexico by the shrimp fishery, which are reported in Table 2.7 of the Assessment Workshop Report, differ from the values in Table 3.10 of the Data Workshop Report. The latter values match those reported in SEDAR-DW06. There is no explanation in the Assessment Workshop Report to explain this inconsistency. Although the Assessment Workshop Report refers to a data workshop report for SEDAR 22 for details of the methods employed to obtain these bycatch estimates, frequent other references to SEDAR 22 in the Assessment Workshop Report suggest that the references to SEDAR 22 are erroneous and that the correct citation should have been the Data Workshop Report for SEDAR 28. This last report provides no explanation for the inconsistency between the values presented in the two reports.

The Assessment Workshop Report presents a table (Table 2.8) of annual standardized estimates of effort for 1945-2011 by the shrimp fishery. These effort values are inconsistent with the effort (days fished) for 1981-2010, which are reported in Table 3 of SEDAR-DW06. While this could possibly have been explained by the fact that the values in Table 2.8 of the Assessment Workshop Report have been standardized, there is no explanation as to how the data for these estimates were collected, nor the method employed to standardize the values. As a further complication, the Assessment Workshop Report advises that the values of effort for the shrimp fishery were input as an index of fishing mortality for the shrimp fishery and. while it would therefore have been expected that the effort values used in the Stock Synthesis model would have been those values reported in Table 2.8 of the Assessment Workshop Report, this is not the case. While there is a broad degree of similarity, the values that are actually input into Stock Synthesis 3, as shown in the data file listed in Appendix A of the Assessment Workshop Report, differ considerably from those presented in Table 2.8. No explanation for this inconsistency is to be found in the cobia Assessment Workshop Report, however the time series of values of effort used in the Stock Synthesis data file for cobia appears to match the time series of scaled effort for the shrimp fishery presented in Table 2.8 of the Assessment Workshop Report for Spanish mackerel. Although this inconsistency thus appears to have a possible explanation, it is important that the results of the stock synthesis runs, estimates of benchmarks, and determinations of current stock status, which have been reported for cobia in the cobia Assessment Workshop Report, are based on the input data for Stock Synthesis that were described in the appendices of that assessment report. Inconsistencies between the data inputs for cobia that have been described and the Stock Synthesis data files for that species need to be reconciled.

The Data Workshop noted that the CVs of the estimates of bycatch of cobia by the shrimp fishery ranged from 66 to 208%, with only 4 of the 39 years having CVs less than 100%. An issue that may have been resolved after the Data Workshop was that a number of the estimates of bycatch calculated by the Bayesian model became stuck on bounds, although the Data Workshop Report does not identify which of the 39 years encountered such problems. As a consequence of these issues, bycatch estimates for the shrimp fishery were recognised by the Assessment Workshop as being very imprecise. For this reason, shrimp fishery effort was used as a proxy for the trends present in the point estimates of bycatch by the shrimp fishery. The median of the 1972 to 2011 estimates of bycatch was used, however, to provide an estimate of the magnitude of the bycatch. An estimate of the catchability coefficient relating shrimp effort to fishing mortality was then calculated within Stock Synthesis using 1972 to 2011 as a super period. A similar super period approach was employed in Stock Synthesis to accommodate the small sample sizes of the length composition data from the SEAMAP program, which were considered to be representative of the length compositions of cobia caught by the shrimp fishery. Use of such a super period to deal with the imprecision of the bycatch estimates of cobia from the shrimp fishery is an appropriate modelling approach. It would have been preferable, however, to have used a reliable time series of precise estimates of discards of the bycatch of cobia from the shrimp fishery in the Stock Synthesis model if such a time series had been available, rather than having to "work around" the problem. Consideration therefore should be given to establishing a well-designed program to monitor the bycatch of cobia by the shrimp fishery such that reliable estimates can be collected in the future.

Very few samples of landed fish were available from catches taken by commercial miscellaneous gears, and thus reliable characterization of the length composition of these landings is not possible, The Data Workshop advised that sample sizes for developing length compositions of commercial landings were inadequate for a considerable number of gears and years. It is reasonable to conclude that length composition data collected from the commercial landings are imprecise. Low sample sizes may also affect the extent to which the resultant length compositions are representative of total annual landings. After filtering, too few measurements of discarded cobia were available from the Reeffish Observer Program to characterize the length composition of discarded fish. The Data Workshop Report advised that age compositions of commercial catches were inadequate for all years and that no aging error matrix could be generated for these ageing data because 86% of the age readings were from a period 15-20 years earlier and thus reader comparisons were not possible. Well-designed monitoring programmes to collect length and age composition data from the landings and discards by each of the principal gear types used by commercial fishers should be established.

Recreational landings

When combining the time series of data collected by different approaches for the same fishing mode, calibration factors were calculated using the data collected during a period of overlap. No comment is made in the Data Workshop Report, but it should be recognised that imprecision of the calculated calibration factor adds to the imprecision of the data that are adjusted and should be carried through into the resulting time series.

While CVs of the estimates of the recreational landings for a fishing mode are calculated and reported in summaries for a number of the data collection programs, estimates of the uncertainty of the values in the resulting time series of the total recreational landings are not provided (Table 2.4, Assessment Workshop Report), and thus are not considered in the assessment.

The collection of age data from the landings of the recreational fishery appears opportunistic, judging from the description provided in the Data Workshop Report. A well-designed program to collect length and age composition data for Gulf of Mexico cobia from the landings and discards of the recreational fishery should be established.

Survey indices

The decisions made by the Data Workshop when selecting indices of abundance appear sound. Despite the fact that both were derived from fishery-dependent data, the time series of headboat and MRFSS catch-per-unit-of-effort (cpue) data were endorsed by the Data Workshop as acceptable indices of abundance for Gulf of Mexico cobia. The time series of data for these indices were standardized using the delta lognormal model.

Adjustment by Assessment Workshop

Although the Data Workshop produced time series of commercial landings by gear type, the Assessment Workshop pooled these data to create a single time series, which was input to Stock Synthesis. Similarly, the Assessment Workshop combined the recreational landings, which had been tabulated by mode, into a single time series of recreational landings. Such pooling obviously suited the incremental approach that was used when developing the assessment model, *i.e.*, first developing a simple production model, then an age-structured production model, and finally a length-structured catch-at-age model. By pooling the data into the two time series, the number of parameters to be estimated was reduced but, as a common selection curve is applied to each time series of combined data within Stock Synthesis, it is assumed that annual length and age-at-length data for the pooled data were representative of those combined data.

ToR 2. Evaluate the quality and applicability of methods used to assess the stock.

Conclusions

Stock Synthesis 3, the software within which the model for the Gulf of Mexico stock of cobia was developed, has gained international recognition for its quality and the applicability of the methods it uses to assess the condition of fish stocks. The model for cobia was of an appropriate structure given the data that were available. Values predicted by the model, including those of benchmarks, were imprecise, however, due to the nature of the input data. Further imprecision of model outputs due to alternative values of key parameters, such as natural mortality and steepness of the stock-recruitment relationship, was explored. Recognising the types of data that were available for input and the uncertainty of model outputs that arose as a consequence of the nature of those input data, the Stock Synthesis model for cobia is of a quality consistent with that which would be considered "best practice", and is able to provide a valuable assessment of the likely condition of the stock in 2011, and, when projected, the likely trajectory of yields and stock condition over the next five to six years.

Strengths

- The decision to use Stock Synthesis 3 as the modelling framework.
- The structure of the model for cobia, which was developed within the Stock Synthesis framework, was appropriate given the data that were available.
- The enhancement of Stock Synthesis to allow modelling of a fishery for which the only source of mortality is that associated with discarding of bycatch.
- The assessment of the uncertainty of parameter estimates was thorough.
- Selectivity runs explored key uncertainties and demonstrated appropriateness of conclusions regarding the current condition of the stock.
- Benchmarks were appropriately calculated.
- Projections were undertaken using two states of nature.

Weaknesses

- Subjective decision to set effective sample size to actual sample size capped at a maximum of 100 rather than to use iterative reweighting, such as proposed by Francis (2011).
- Lack of exploration of sensitivity to the assumption of logistic selectivity for the recreational and commercial fisheries.
- Lack of length and age composition data to provide information on the length compositions of discards and the shape of the retention curves
- Failure of model to match the trends in discards from the commercial and recreational fisheries
- Imprecision in the estimate of steepness of the stock-recruitment relationship.
- Lack of exploration of uncertainty associated with time series of commercial and recreational landings.
- Errors in Stock Synthesis files in the Appendices.

Both the decision by the Assessment Workshop to employ Stock Synthesis 3 as the modelling framework and the structure of the model for the Gulf of Mexico stock of cobia that was developed within this framework are appropriate. Stock Synthesis has been extensively tested, and has the flexibility to be applied to fisheries with data qualities ranging from poor to rich. The software has been equipped with tools to explore uncertainty, to estimate benchmarks, and to undertake projections using alternative harvest policies. Because of its versatility, Stock Synthesis is well suited to explorations of the sensitivities of model outputs to a broad range of alternative model structures or use of alternative sets of data inputs. The enhancement of Stock Synthesis to allow modelling of a fishery for which the only source of mortality is that associated with discarding of bycatch is a particular strength of the assessment that was developed for the Gulf of Mexico stock of cobia. While some deficiencies were identified in the fit of the base model, the overall fit was regarded as adequate.

The Stock Synthesis model for the Gulf of Mexico stock of cobia included three fishing fleets, *i.e.*, commercial, recreational and discards of bycatch from the shrimp trawl fishery, and two fishery-dependent abundance indices, *i.e.*, cpue data from the MRFSS survey and from the headboat survey. Time series of discards from the commercial, recreational, and shrimp fisheries were input, together with length composition data of cobia from the commercial and recreational fisheries, and, combining the data into a super period, from the bycatch from the shrimp fishery. Age composition data were input for the recreational fishery and considered within the model as age compositions that were conditional on length.

The model employed 3-cm bins for the length composition of cobia, and the lower bounds of the length intervals within these bins ranged from 6 to 165 cm. It was pleasing to note that the Assessment Workshop had reported exploration of the effect of bin size on estimation of selectivity parameters, at least to a limited extent, and concluded that use of a bin width of 3 cm was preferable to use of one that was 5 cm. Methot (2011) notes that, on occasion, wide bin widths can cause problems when the slope of a selectivity or retention curve becomes so steep that all change occurs within a single length class.

Although the Assessment Workshop reported that, as its value is typically unable to be estimated within the assessment model, the standard deviation of recruitment was fixed at 0.6, no justification for the choice of this particular value is provided in the Assessment Workshop Report. It might be useful to note that the use of this value has been proposed in a number of studies (*e.g.*, Smith and Punt 1998; Maunder and Deriso, 2003), which typically advise that the value 0.6 is supported by the results of the meta-analyses undertaken by Beddington and Cooke (1983), and later by Mertz and Myers (1996).

When developing the base model for cobia, a subjective decision was made to employ an effective sample size for the length composition data of cobia, which was set equal to annual sample size but capped at a maximum of 100 when the number of fish in the annual sample exceeded this number. Rather than using this subjective approach, the iterative re-weighting approach that was explored in sensitivity run 10, *i.e.*, the method proposed by Francis (2011), is recommended.

The decisions by the Assessment Panel to use asymptotic, logistic, size-based selectivity curves for the recreational and commercial fisheries and a double-normal selectivity curve to represent the selectivity of cobia by the shrimp fishery, and to keep these selectivity curves constant over time, are endorsed. It would have been expected, however, that sensitivity to this choice of selectivity patterns would have been explored. As was appropriate, to accommodate the introduction in 1984 of a minimum size limit of 33 inches, separate retention curves were assumed for the time blocks 1927-1984 and 1985-2011. Because of the lack of data prior to 1993, however, it was necessary to assume the shape and parameters of the retention curve for the earlier time block. This represents a source of uncertainty, and it would therefore be appropriate to consider whether assessment results are likely to be sensitive to the assumptions made regarding the form and values of parameters of this retention curve.

The base model was fitted to the data for Gulf of Mexico stock of cobia and reported as Run 1. All estimated parameters were assumed to have uniform, non-informative priors, with wide bounds. The results of the jitter test, with 48 of 50 trials converging to within 2 likelihood units of the minimum, suggested that the model was not particularly sensitive to the initial values of the parameters that were estimated.

While model predictions were broadly consistent with the commercial and recreational discards, the trends of the predictions did not match those of the observed data, suggesting some structural deficiency of the model or, if the model structure was correct, inadequacy of the discard data or overriding influence of other data. In the case of discards by the commercial fishery, the possibility that the discard data were inadequate cannot be discounted as the Data Workshop had identified that these estimates were likely to be both imprecise, as few trips with cobia discards had been recorded in the Reeffish Observer Program, and erroneously low, as the NMFS logbooks do not provide coverage of the entire fishery. In the case of the recreational fishery, however, it is likely that the failure to fit the trend in recreational discards was due to the competing influence of other datasets on model predictions.

It would be useful to advise in the captions of Figures 3.7 and 3.8 of the Assessment Workshop Report that these are plots of the MRFSS and headboat cpue data, respectively. As noted in the Report, the fits to these indices and to the effort data for the shrimp fishery are quite good, although runs of positive and negative deviations were present in the headboat cpue data. Some structure also appeared present in the Pearson residual plots for the commercial (Fig. 3.11) and recreational (Fig. 3.13) length composition data.

In the base model represented by Run 1, estimates of both the log of unexploited equilibrium recruitment (1,033,130 fish) and the steepness of the stock recruitment curve, *i.e.*, 0.925, were calculated by Stock Synthesis when the model was fitted to the input data. The Assessment Workshop provided a well-considered evaluation of the reliability of the estimate of steepness, noting that a large proportion of bootstrap estimates of steepness approached the upper bound of 1, and that, although probably greater than 0.8, the distribution of estimates between 0.85 and 1 was relatively uniform. The likelihood profile for steepness was relatively flat between 0.8 and 1, but suggested a minimum between 0.85 and 0.95. Tension was exhibited in the values of steepness that were most consistent with recruitment data (favouring a value of ~1), length and discard data (favouring a value of ~0.8), and age composition (favouring a value of ~0.65), with little information relating to steepness evident in the abundance indices. The fact that the input data were more consistent with lower values of steepness, while the assumption regarding recruitment deviations appeared to be providing the support for higher values of steepness, is interesting as it raises the question of whether, in the case of Gulf of Mexico cobia, the influence of recruitment deviations on the resultant parameter estimates was excessive. The assessment Workshop Report advised that steepness may not be well estimated by the Stock Synthesis model, a conclusion that appears sound. The recent study by Lee *et al.* (2012), which demonstrated the difficulty that is typically encountered when attempting to estimate steepness, concluded that "steepness is reliably estimable inside the stock assessment model only when the model is correctly specified for relatively low productive stocks with good contrast in spawning biomass". This conclusion is relevant to the cobia assessment, for which the results of fitting the base model to cobia, a species that, on the basis of its natural mortality, would be considered of medium productivity, indicated that biomass had been relatively stable over the last 30 years, the period covered by the abundance indices and much of the more reliable input data.

The question of how to respond when the steepness of the stock-recruitment relationship is imprecise or cannot be estimated reliably should be considered. Francis (2012) has suggested that, in such circumstances, he considers it better to fix steepness at a value, such as 0.75, *i.e*., the default value recommended in Francis (1993), and which is frequently used in Australia and New Zealand, or the average of published values for the same or similar species. Francis (2012) advises that the uncertainty associated with this parameter should then be explored using sensitivity runs with lower and higher values of steepness.

There would have been value in assessing whether the value of steepness estimated from the base model, *i.e.*, 0.925, is consistent with published values for cobia or similar species. The fact that this value of steepness for the base model, and the values of steepness estimated when fitting the models using the low and high values of the base level of natural mortality, which were subsequently used as alternative states of nature, ranged from 0.92 to 0.96 (Table 3.7, Assessment Workshop Report) was initially of concern to the Reviewer, as such values of steepness reflect a robust stock that is able to maintain recruitment despite considerable decline in stock size. It was noted subsequently, however, that the Assessment Workshop had explored sensitivity runs with lower steepness, *i.e*., 0.7 and 0.8, and that these runs had produced very similar conclusions regarding the condition of the stock with respect to benchmark levels as were determined using the base model (Table 3.8, Stock Assessment Report). Accordingly, after considering the results of the other sensitivity runs, it is concluded that, despite imprecision in the estimate of steepness, the base model accepted by the Assessment Workshop, *i.e*., the model associated with Run 1, is appropriate for determination of the current condition of the Gulf of Mexico stock of cobia and for use in projecting the fishery over a short time period to assess the likely outcomes of fishing with specified levels of fishing mortality.

There are errors in the stock synthesis files listed in the appendices. For example, there are actually 91 length observations in the data file, not 85, where this inconsistency would cause Stock Synthesis to abort when it attempted to read the data. Also, the number of length bins is specified as 54 in the data file, but the specification of the selectivity for MRFSS data attempts to use 57, which would cause Stock Synthesis to abort when it attempted to run following data input. The listings should be those associated with the base model, but appear to be those of a model that was still under development.

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

Conclusions

Estimates of stock abundance, biomass, and exploitation are produced when the Stock Synthesis model is fitted. The values of total biomass and annual exploitation in 2011, which were estimated when the base model for the Gulf of Mexico stock of cobia was fitted, were 3,030 mt and 0.29, respectively.

Strengths

Stock Synthesis 3 is able to calculate time series of abundance, total biomass, and annual exploitation.

Stock abundance:

The report file that is produced by Stock Synthesis, report.sso, contains a time series section, in which the time series of abundance, recruitment and catch for each of the areas are reported. Output quantities include summary biomass and summary numbers for each gender and growth pattern. The Assessment Workshop Report for the Gulf of Mexico cobia stock has not reported these abundance estimates, but they will be available in the output file for Run 1.

Biomass:

Stock Synthesis produces an estimate of total annual biomass (Table 3.4, Fig. 3.33). The estimate (for Run 1) of total biomass for 2011 was 3,030 mt.

Exploitation:

Although not reported in the text of the Assessment Workshop Report, the code within the Starter.SS file presented in Appendix C of this report specifies that, for the Gulf of Mexico stock of cobia, Stock synthesis is to set the value of fishing mortality, F, to the value of annual exploitation, calculated as the ratio of the weight of the total catch (including discards) to the total biomass. The estimate (for Run 1) of the annual exploitation rate for 2011 was 0.29 (Table 3.6, Assessment Workshop Report).

ToR 4. Evaluate the methods used to estimate population benchmarks and management parameters. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.

Conclusion

Stock Synthesis calculates a range of population benchmarks and management parameters. Benchmarks calculated for cobia were MFMT = $F_{30\%\text{SPR}}$ and MSST = (1 – M) $\text{SSB}_{30\%\text{SPR}}$. The estimates of F_{current} and $\text{SSB}_{\text{current}}$, which were calculated for 2011 using the base model for cobia, were 0.24 and 2,213 mt, respectively. The ratios $F_{\text{current}}/\text{MFMT}$ and $\text{SSB}_{\text{current}}/\text{MSST}$, which were calculated using the base model, were 0.63 and 1.73, respectively. These results, which were consistent with those produced by all but one (the model with natural mortality set to 0.26 y^{-1}) of the models used in the various sensitivity runs, imply that, in 2011, the Gulf of Mexico stock of cobia was not experiencing overfishing and was not overfished.

Strengths

Stock Synthesis possesses well-tested procedures to calculate and output a range of population benchmarks and management parameters.

Summary

Stock Synthesis provides estimates of population benchmarks and management parameters. In particular, it is able to produce estimates for indicator variables and reference points based on maximum sustainable yield (MSY), spawning potential ratio (SPR), and spawning stock biomass (SSB), and taking the stock-recruitment relationship into account. SPR is calculated as the equilibrium spawning biomass per recruit that would result from a given year's pattern and the levels of *F*'s and selectivities for that year. For MSY-based reference points, Stock Synthesis searches for a fishing mortality that would maximise the equilibrium yield. For SPR-based reference points, the computer program searches for an *F* that would produce the specified level of SPR. For spawning biomass-based reference points, the software searches for an *F* that would produce the specified level of spawning biomass relative to the unfished value.

The management benchmarks, *i.e.*, the Maximum Fishing Mortality Threshold (MFMT) and Minimum Stock Size Threshold (MSST), which were proposed for the Gulf of Mexico stock of cobia by the Assessment Workshop, are appropriate for use in determining the status of that stock. These benchmarks, which were based on the level of fishing mortality and equilibrium spawning stock biomass associated with a spawning potential ratio of 30%, are

$$\text{MFMT} = F_{30\%\text{SPR}} \quad \text{and} \quad \text{MSST} = (1 - M)\ \text{SSB}_{30\%\text{SPR}},$$

where it was concluded that overfishing was occurring if F_{current} > MFMT, *i.e.*, $F_{\text{current}}/\text{MFMT} > 1$, and the stock was considered to be overfished if $\text{SSB}_{\text{current}}$ < MSST, *i.e.*, $\text{SSB}_{\text{current}}/\text{MSST} < 1$.These benchmarks are approximations for

$$\text{MFMT} = F_{\text{MSY}} \quad \text{and} \quad \text{MSST} = (1 - M)\ \text{SSB}_{\text{MSY}},$$

where F_{MSY} is the fishing mortality that produces the maximum sustainable yield MSY, M is the point estimate of natural mortality for fully recruited ages, and SSB_{MSY} is the equilibrium spawning stock biomass that produces MSY. The benchmarks for the Gulf of Mexico stock of cobia use proxies, where these proxies were based on a spawning potential ratio SPR of 30%. Thus, the proxy that was used for F_{MSY} was the fishing mortality, $F_{30\%\text{SPR}}$, which produces a spawning stock biomass per recruit that is 30% of the spawning stock biomass per recruit produced when the stock is not fished, *i.e.* an SPR of 30%. The proxy that was used for SSB_{MSY} was the corresponding value of equilibrium spawning stock biomass, *i.e.* the spawning stock biomass $\text{SSB}_{30\%\text{SPR}}$ that is produced with a fishing mortality of $F_{30\%\text{SPR}}$.

Although Stock Synthesis is able to estimate MSY-based rather than SPR-based reference points, the Assessment Panel chose to use the proxies $F_{30\%\text{SPR}}$ and $\text{SSB}_{30\%\text{SPR}}$ rather than F_{MSY} and SSB_{MSY}. The latter two reference points are likely to be more appropriate if assessing "the capacity of a fishery to produce the maximum sustainable yield on a continuing basis" (Magnuson-Stevens Fishery Conservation and Management Act, May 2007).

F_{current} was calculated as the geometric mean of the estimates of the three most recent annual fishing mortalities, *i.e.,* the fishing mortalities for 2009-2011, where annual fishing mortality was estimated by its proxy, exploitation rate, calculated as the ratio of the total catch (including discards) to estimated total biomass. $\text{SSB}_{\text{current}}$ was the estimate of spawning stock biomass for 2011.

Table 3.8 of the Assessment Workshop Report, a subset of which is reproduced below, contains the values of the current (2011) fishing mortality and spawning stock biomass for Gulf of Mexico cobia, the values of the MFMT and MSST benchmarks for this stock, and the results of the stock determination for each of the models that were explored in the assessment. The only one of these models, for which the current fishing mortality exceeded MFMT (*i.e.*, overfishing was occurring) or the current SSB was less than MSST (*i.e.*, the stock was overfished), was the sensitivity trial in which a low value of natural mortality was employed as the base level when scaling the Lorenzen (1996) estimates to determine age-dependent estimates of natural mortality.

Quoted from Assessment Workshop Report: "Table 3.8, Assessment Workshop Report. Reference points and benchmarks from sensitivity runs for Gulf of Mexico cobia from SS. Benchmarks are reported for SPR 30%. Current refers to the geometric mean of 2009-2011 for F. MSST = (1-M)*SSBSPR30% with M = 0.38 y^{-1} for all models except runs 2 (M = 0.26 y^{-1}) and 3 (M = 0.50 y^{-1})".

Run	Model	Fcurrent	SSB2011	MFMT	MSST	F/MFMT	SSB/MSST	Overfishing occurring?	Overfished?
1	Base model	0.24	2213	0.38	1280	0.63	1.73	No	No
2	M_Low	0.3	1872	0.29	2443	**1.05**	**0.77**	**Yes**	**Yes**
3	M_High	0.18	2587	0.45	804	0.4	3.22	No	No
4	D_High	0.24	2197	0.37	1302	0.65	1.69	No	No
5	Steepness=0.7	0.24	2121	0.39	1174	0.63	1.81	No	No
6	Steepness=0.8	0.24	2168	0.38	1257	0.64	1.73	No	No
7	MRFSS only	0.26	1921	0.37	1277	0.7	1.5	No	No
8	HB only	0.19	2940	0.37	1301	0.52	2.26	No	No
9	Stock synthesis weighted	0.22	2340	0.35	1273	0.58	1.85	No	No
10	Francis (2011) weighting	0.22	2415	0.38	1305	0.61	1.84	No	No

ToR 5. Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.

Conclusions

Stock Synthesis provides a well-tested procedure to project the model through a range of future years, using a fishing rate based on MSY, SPR, a specified target biomass, or a multiple of the recent average fishing rate, and producing estimates of yield and key management parameters, thereby allowing assessment of future stock condition. The methods used, which are recognised as being of high quality, are designed to produce the estimates of future population status that are needed by managers. For the base model, fishing mortality would be increased from F_{current} if adjusted to F_{OY} or $F_{30\%\text{SPR}}$. Projections from 2013 to 2019 suggest that spawning stock biomass would increase from $\text{SSB}_{\text{current}}$ if fishing mortality was maintained at F_{current}, increase to a lesser extent if fishing mortality was increased to F_{OY}, and decline very slightly if fishing mortality was increased to $F_{30\%\text{SPR}}$. Yield would be expected to increase under each of these three fishing mortalities. The condition of the stock would be expected to continue to be classified as "not overfished, with overfishing not occurring".

Strengths

Projections are undertaken using the well-tested procedures within Stock Synthesis.

Weaknesses

It would have been useful to have undertaken a projection using a model with a lower steepness, such as 0.8.

Summary

Stock Synthesis includes a well-tested procedure to project the future stock status that would be expected to result when using a fishing rate based on MSY, SPR, a specified target biomass, or a multiple of the recent average fishing rate. Use of this procedure ensures consistency of model predictions with the assumptions, with the parameter estimates obtained by fitting the model, and with the length and age structure predicted as the current state of the stock. It is thus highly applicable for use with the Gulf of Mexico stock of cobia.

Deterministic projections for 2013 to 2019 were run for the Gulf of Mexico stock of cobia using three models, *i.e.*, the base model (Run 1), and the low and high mortality models (Runs 2 and 3, respectively), which the Assessment Panel considered representative of possible alternative states of nature. The projections were made using fishing rates set to MFMT (*i.e.*, the proxy $F_{30\%SPR}$ for F_{MSY}), F_{OY} (*i.e.*, 75% of $F_{30\%SPR}$), and $F_{current}$, where this last value was calculated as the geometric mean of the annual values of F for the last three years, *i.e.*, 2009-2011. The fishing mortality of the shrimp fishery during the projection period was assumed to remain constant, and was set to the geometric mean of the annual fishing mortalities for this fishery over the last three years, *i.e.*, 2009-2011. Selectivity, discarding, and retention patterns were assumed to be the same as those experienced in the five most recent years, *i.e.*, 2007-2011, while the distribution of catches among the fishing fleets, *i.e.,* fisheries, reflected the distribution of average fishing intensities among those fleets in 2009-2011. Recruitment during the projection period was calculated as the value predicted by the stock-recruitment relationship. The base model was also projected using a fishing mortality of $F_{30\%SPR}$ for 1000 samples generated using the bootstrap facility within Stock Synthesis to produce distributions of the estimated yields predicted by the model for each year between 2012 and 2019 (Fig. 3.63, Assessment Workshop Report).

The final year of the time series of data used in the assessment for the Gulf of Mexico stock of cobia was 2011. In order to carry out projections, it was therefore necessary to estimate the removals that were likely to have occurred in 2012. Accordingly, removals of cobia for each of the fisheries in 2012 were estimated using a fixed fishing mortality set to the geometric average of the annual fishing mortalities in 2009-2011.

The methods used in Stock Synthesis to predict the outcomes expected between 2013 and 2019 were considered to be of a high quality. The quality of the resulting projections depends, however, on the extent to which the alternative states of nature represented by the different models used in the projection are likely to be representative of the true state of nature, and the extent to which each of those alternative models provides a reliable representation of the dynamics of the stock. The results of the projections should thus be considered in the context of the accuracy and precision of the predictions made by the model with respect to the input data they were intended to represent.

Although the three models used in the projections bracket the range of estimates of natural mortality for cobia, the estimates of steepness for these models range only between 0.92 and 0.96, *i.e.*, there will be little reduction in recruitment as spawning stock biomass declines, until the depletion in spawning stock biomass becomes severe. There would have been value in considering a model with a considerably lower value of steepness, *e.g.*, 0.8, to represent an alternative state of nature, which, given the nature of the input data and the uncertainty of the estimate of steepness, appears feasible.

The results obtained from the projections are presented in Table 3.9 and Figures 3.59-3.70 of the Assessment Workshop Report. Estimates of stock condition depend on which of the states of nature explored in the assessment is most likely to reflect the true state of nature. Of the three scenarios considered in the assessment, that represented by the base model (Run 1) would be considered to provide the best description of the data that were available, given the assumptions that were made regarding those data, the biology of the cobia stock, and the fisheries exploiting this stock. For the base model, fishing mortality would be increased from F_{current} if adjusted to F_{OY} or $F_{30\%\text{SPR}}$. The base model predicts that spawning stock biomass would be expected to increase from $\text{SSB}_{\text{current}}$ if fishing mortality was maintained at F_{current}, increase to a lesser extent if fishing mortality was increased to F_{OY}, and decline very slightly if fishing mortality was increased to $F_{30\%\text{SPR}}$. Yield would increase under each of these three fishing mortalities. If the model with the lower natural mortality, *i.e.*, Run 2, represented the true state of nature, continued fishing with a fishing mortality of F_{current} is predicted to allow the spawning biomass to increase beyond the MSST by 2014, *i.e.*, become no longer overfished, despite the fact that overfishing was continuing. The reduction in fishing mortality associated with F_{OY} or $F_{30\%\text{SPR}}$ would result in overfishing no longer occurring and would produce an increase in spawning stock biomass such that, by 2014, the stock would no longer be classified as being overfished. If natural mortality was greater, *i.e.*, Run 3, spawning stock biomass would increase if fishing mortality was maintained at F_{current} but would decline if it was set to F_{OY}, and would decline to an even greater extent if fishing mortality was set to $F_{30\%\text{SPR}}$.

It would have been informative to explore the consequences (for each pair of putative states of nature) of incorrectly assuming that one of these alternative states of nature was true, and setting allowable catches accordingly, when in fact one of the alternative states of nature was the "true" state. Such an analysis allows an assessment of the robustness of an incorrect decision relating to which of the alternative models is considered most likely to represent the true state of nature.

ToR 6. Evaluate the quality and applicability 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**
- **If there are significant changes to the base model, or to the choice of alternate states of nature, then provide a probability distribution function for the base model, or a combination of models that represent alternative states of nature, presented for review.**
- **Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments**
- **Provide justification for the weightings used in producing the combinations of models**

Conclusions

The methods within Stock Synthesis that may be used to explore uncertainty include calculation of estimates of asymptotic standard errors, calculation of likelihood profiles, MCMC analyses, and bootstrapping. These tools are complemented by auxiliary routines that allow production of diagnostic plots, which also assist in communicating the uncertainty of estimates. The software encourages exploration of alternative model structures and sensitivity to alternative values of parameters or functional forms. The model that was developed for the Gulf of Mexico stock of cobia employed an appropriate set of these methods. Probability distributions were produced for initial equilibrium biomass and steepness, unfished total and spawning biomass, and spawning biomass in 2011. As the iterative approach required to calculate *P** cannot be implemented in Stock Synthesis, Stock Synthesis "calculates the expected time series of probabilities that the *F* resulting from a specified harvest policy would exceed a specified level" (Methot and Wetzel, 2012).

Strengths

- Stock Synthesis provides an extensive suite of methods that may be used to explore uncertainty.
- The retrospective analysis revealed no strong systematic patterns.
- Bootstrapping was used to produce probability distributions

Summary

Stock Synthesis provides a number of methods that may be used to characterize the uncertainty associated with the estimates of parameters, benchmark estimates, and predicted values of parameters. These include options to generate likelihood profiles and to run a bootstrapping or Markov Chain Monte Carlo (MCMC) analysis. The software is well suited for use in exploring the uncertainty associated with the models that were fitted to the Gulf of Mexico stock of cobia. Thus, for each run of the Stock Synthesis model for this stock, estimates of asymptotic standard errors would have been calculated for each of the parameters that were estimated (see Table 3.1, Assessment Workshop Report, for parameter estimates and estimates of asymptotic standard errors for the base model, Run 1, for which the average value of natural mortality for fully-selected cobia was $M = 0.38\ y^{-1}$ and estimated steepness = 0.925). These standard errors may be considered to represent minimum values for the uncertainty of the estimated parameters. The uncertainty of selected parameter estimates for the Gulf of Mexico cobia stock was also characterized using the results from bootstrapping (Table 3.2, Figs 3.26 and 3.27). Additional uncertainties (sensitivities) arising from differences in model structure or data input for the cobia model were also assessed by re-running Stock Synthesis using those alternative model structures or data sets.

The initial run (Run 1) was carried out using the model structure that had been proposed for the Gulf of Mexico stock of cobia and estimating the steepness parameter of the Beverton and Holt stock-recruitment relationship. Bootstrapping of this model demonstrated that, given the data that were available, the steepness of the stock recruitment relationship was estimated imprecisely, a result which was confirmed by constructing likelihood profiles for this parameter. A number of sensitivity runs of Stock Synthesis were then run to explore the effect of varying this and other parameters, or the methods employed in the analysis.

As is typical in stock assessment, exploratory runs for the Gulf of Mexico stock of cobia were first employed to determine a base model for the assessment, *i.e.*, a model that is considered the most likely of the alternative model configurations that have been proposed. Despite the imprecision of the estimate of steepness, the decision was made at the Assessment Workshop to retain Run 1 as the base model as parameter estimates and patterns of stock dynamics were similar for the models using alterative estimates of steepness.

The Assessment Workshop selected the models with low *M* (Run 2) and high *M* (Run 3) as representative of alternative states of nature. Projections using these models were explored.

While the iterative approach required to calculate *P** cannot be implemented in Stock Synthesis, a complementary approach has been developed to produce estimates of the probability that *F*, the fishing rate based on MSY, SPR, a specified target biomass, or a multiple of the recent average fishing rate that is employed in the projection, exceeds the OFL (Methot and Wetzel, 2012). These authors advise that, whereas the *P** approach calculates the future stream of annual catches that would have a specified annual probability of F > OFL, Stock Synthesis "calculates the expected time series of probabilities that the *F* resulting from a specified harvest policy would exceed a specified level".

The models were not combined, but presented as alternatives for consideration by the Review Panel.

ToR 7. If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.

The Review was undertaken as a desktop review, rather than a review within a workshop setting. Accordingly, it was not possible for the recommendations made in review reports to be acted upon, nor to ensure that the results were incorporated accurately in the resultant Stock Assessment Report.

ToR 8. Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.

The SEDAR Process provides a very sound basis for stock assessment. It has ensured that all aspects of the assessment process for the Gulf of Mexico cobia, from collation of data through to model development, exploration, and production of management advice, have been documented in detail, including the underlying reasons for decisions that were made concerning data to be used and model structure to be employed. For the reviewer, it has thus provided a thorough understanding of the details of the assessment and assisted in identifying opportunities for improvement and in detecting errors or inadequacies.

The Terms of Reference for the Assessment Process, which are presented below, are now examined and comment is made on the degree to which these were addressed.

1. Review and provide justifications for any changes in data following the data workshop and any analyses suggested by the data workshop. Summarize data as used in each assessment model.

   Accomplished.

2. Recommend a model configuration which is deemed most reliable for providing management advice using available compatible data. Document all input data, assumptions, and equations.

   The configuration of the model for cobia that was set up within the Stock Synthesis framework was described. The equations used within Stock Synthesis were not described in the Assessment Workshop Report. This is understandable as, to some extent, the rate of development of this software has outpaced the development of the technical descriptions relating to the features within the Stock Synthesis software. Methot and Wetzel (2012) have recently addressed this issue, however, and their recent paper should be cited in the Assessment Workshop Report.

3. Incorporate known applicable environmental covariates into the selected model, and provide justification for why any of those covariates cannot be included at the time of the assessment.

   No environmental covariates were identified by the Data or Assessment Workshops.

4. Provide estimates of stock population parameters.
   • Include fishing mortality, abundance, biomass, selectivity, stock-recruitment relationship, and other parameters as appropriate given data availability and modeling approaches
   • Include appropriate and representative measures of precision for parameter estimates

   Accomplished.

5. Characterize uncertainty in the assessment and estimated values.
   • Consider components such as input data, modeling approach, and model configuration
   • Provide appropriate measures of model performance, reliability, and 'goodness of fit'

   Accomplished.

6. Provide yield-per-recruit, spawner-per-recruit, and stock-recruitment evaluations.

   Accomplished.

7. Provide estimates of stock status relative to management criteria consistent with applicable FMPs, proposed FMPs and Amendments, other ongoing or proposed management programs, and National Standards for each model run presented for review.

   Accomplished.

8. Project future stock conditions and develop rebuilding schedules if warranted, including estimated generation time. Develop stock projections in accordance with the following:
   A) If stock is overfished:
      F=0, FCurrent, FMSY, FOY
      F=FRebuild (max that permits rebuild in allowed time)
   B) If stock is undergoing overfishing:
      F= FCurrent, FMSY, FOY
   C) If stock is neither overfished nor undergoing overfishing:
      F= FCurrent, FMSY, FOY
   D) If data limitations preclude classic projections (i.e. A, B, C above), explore alternate models to provide management advice

   Accomplished.

9. Provide a probability distribution function for the base model, or a combination of models that represent alternate states of nature, presented for review.
   - Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments for use with the Tier 1 ABC control rule
   - Provide justification for the weightings used in producing combinations of models

   The Assessment Workshop Report noted that three of the sensitivity runs had been considered as alternate states of nature, and projections had been run for each of these The Assessment Workshop Report advised that probability distribution functions had been developed for the subset of three runs and would "be made available to the Scientific and Statistical Committee (SSC) for the development of management advice, including OFL and ABC". No information relating to these probability distribution functions was presented in the Report.

10. Provide recommendations for future research and data collection. Be as specific as possible in describing sampling design and intensity, and emphasize items which will improve assessment capabilities and reliability. Recommend the interval and type for the next assessment.

    Attention was directed to the research recommendations that were made in the Data Workshop Report. The Workshop Assessment Report identified gaps in data, which, if addressed, would improve the assessment capabilities and reliability. Specific sampling design and intensity were not discussed. No recommendations relating to the interval and type for the next assessment were made by the Assessment Workshop.

11. Prepare a spreadsheet containing all model parameter estimates and all relevant population information resulting from model estimates and projection and simulation exercises. Include all data included in assessment report tables and all data that support assessment workshop figures.

    A spreadsheet was not provided in the documentation that was circulated to the Review Panel. The Assessment Workshop addressed this Term of Reference in its Report by providing a table listing the estimates for all parameters used in the model and presenting a listing of each of the input files required to run the Stock Synthesis model for Gulf of Mexico cobia.

12. Complete the Assessment Workshop Report (Section III: SEDAR Stock Assessment Report).

    Accomplished.

ToR 9. Make any additional recommendations or prioritizations warranted.

- **Clearly denote research and monitoring needs that could improve the reliability of future assessments**

A number of research needs, which are listed below in order of priority, were identified in the course of the desk review. As expected, these were highly consistent with, and thus overlap, many of the research needs that had been identified by the Data and Assessment workshops.

1. Review or establish programs to collect data on the length composition and age-at-length compositions of landings and discards from each commercial gear and from each recreational fishing mode, and of bycatch of cobia from the shrimp fishery. Ensure that the statistical design and spatial coverage of survey or sampling programs are appropriate and that survey or sampling intensity is sufficient to produce estimates of the required precision for Gulf of Mexico cobia. Set goals for performance and establish and monitor performance criteria to assess the quality and completeness of data collection programs. This item is of the highest priority as it will provide information required by Stock Synthesis to determine the selectivity and retention curves for cobia for the commercial, recreational, and shrimp fisheries, the lack of which is a key source of uncertainty in the model.
2. Undertake research to determine reliable relationships between the proportion of females that are mature and both length and age for the Gulf of Mexico stock of cobia. This item is also of high priority, as the maturity information that is currently used is imprecise. The calculation of spawning stock biomass, a crucial parameter in the calculation of benchmarks and assessment of stock status, should be based on reliable data.
3. Review programs that are used to collect discard data for cobia (and data on the bycatch of cobia by the shrimp fishery), and refine these programs to ensure that accurate and complete data estimates of the discards (and bycatch) are collected. Ensure that the statistical design and spatial coverage of survey or sampling programs are appropriate

and that survey or sampling intensity is sufficient to produce estimates for Gulf of Mexico cobia that are of the required precision. Set goals for performance and establish and monitor performance criteria to assess the quality and completeness of data collection programs and provide feedback regarding performance to those programs. While this research item will not provide immediate improvement in the quality of the assessment, it is important that action is taken as soon as possible to improve the accuracy and precision of the data relating to the quantities of fish that are discarded from each of the fisheries, such that, in the future, the time series of discards become more reliable.

4. A comprehensive genetic study of cobia should be undertaken, with the following objectives:
    a. to confirm the preliminary genetic findings of Darden for cobia in the Gulf of Mexico and US Atlantic Coast, using samples with sample sizes greater than 100 at all sites, thereby addressing the issue in that earlier study that sizes of samples from the north of the Gulf of Mexico and from waters off the west coast of Florida had been small;
    b. to increase the spatial resolution of the genetic sampling in the region of overlap of the two stocks, such that the boundary between the stocks or extent of overlap can be determined;
    c. to extend sampling into Mexican waters and thereby determine the southern boundary of the Gulf of Mexico stock;
    d. to reconcile the differences in the findings reported in SEDAR28-DW01 and those reported in SEDAR28-RD09, where the former advises that collections from offshore in the Gulf of Mexico were genetically distinct from those offshore in the South Atlantic region while the latter reports that the results of the study "suggest the offshore groups are genetically homogenous, even between the SA and GOM";
    e. to extend sampling beyond the spawning season and ascertain whether catches of fish may be assigned reliably to either the Gulf of Mexico or South Atlantic stock on the basis of the area in which they are caught.

    Some of the objectives of this study, *e.g.*, identification of the southern boundary of the stock, would also benefit from tagging or other studies. As this study will take some time before completion, it has been assigned a lower priority than the previous items. Determination of the southern stock boundary, however, is important to ensure that other removals from the stock are not occurring in Mexican waters, as such removals are not taken into account in the current assessment.
5. Undertake research to determine the discard mortality of Gulf of Mexico cobia that are discarded from the catches of each commercial fishing gear or each recreational fishing mode, recognising that such mortality is likely to differ among different categories into which the discarded fish are classified, *e.g.*, "alive", "mostly alive", and "mostly dead".
6. In future stock assessments for the Gulf of Mexico stock of cobia, explore whether the use of an age-dependent rather than constant *M* results in a significant improvement in fit, considering the Lorenzen and alternative functional forms of the relationship with age and the alternative of estimating the value of the age-dependent *M* at each age (or range of ages).

7. In future stock assessments, explore the sensitivity of the model to the uncertainty of the landings data.
8. Develop an ageing error matrix for Gulf of Mexico cobia.
9. A research study should be undertaken to determine an approach (or approaches) by which an appropriate range (or ranges) of feasible values of *M* for a species might be selected for use in stock assessment as alternate plausible states of nature. The need to determine an appropriate range for sensitivity runs arose in both the cobia and Spanish mackerel assessments, but the final decisions on the range to use were rather arbitrary and subjective. The issue arises in almost all assessments and it would be useful to establish an objective protocol to determine an appropriate range of values of *M* to be explored.
10. Develop a fishery-independent survey for Gulf of Mexico cobia, or investigate what changes would be required to make data from an existing fishery-independent survey appropriate for use as an index of abundance.
11. As a low research priority, assess whether, in future refinement of the Stock Synthesis model, sexually dimorphic growth should be introduced. Note that the benefit of this might only be realised if appropriate sex composition data for landings and discards are available for input, and length and age-at-length compositions are sexually disaggregated.

*4.2 Gulf of Mexico Spanish Mackerel (***Scomberomorus maculates***)*

ToR 10. Evaluate the quality and applicability of data used in the assessment.

Conclusions

The data compiled for the Gulf of Mexico stock of Spanish mackerel by the Data Workshop are the best that are available. Certainly, some aspects of the data are imprecise, *e.g.*, discards from commercial catches, and there are data gaps, such as the lack of length and age-at-length composition data for discards. Nevertheless, the data that are available are of a quality that would allow a broad assessment of the likely condition of the stock, which, although uncertain, would be useful to fisheries managers.

Strengths

- The collation of life history data for the Gulf of Mexico stock of Spanish mackerel.
- The collation of commercial landings data to produce time series of landings by gillnet, handline, and other gears from 1887, and, particularly, more precise data from 1950.
- The collation of a time series of estimates of bycatch of Spanish mackerel by the shrimp fishery from 1972, using a Bayesian model.
- The collation of recreational fisheries data from different sources to produce sound time series of landings by fishing mode from 1955, and, particularly, more precise data from 1981.
- The collation of data to produce time series of discards from the commercial gears and recreational fishing modes.

- The collation of length composition data to characterize the landings by the commercial and recreational fisheries.
- The collation of a fishery-independent and two fishery-dependent indices of abundance, and the use of appropriate statistical analyses to standardize those indices of abundance.

Weaknesses

- Lack of definition of the southern boundary of the Gulf of Mexico stock of Spanish mackerel.
- Uncertainty of the age at which 50% of Spanish mackerel are mature.
- The unreliable nature of the discard data due to low reporting, low intercept rates, and inadequate data collection programs.
- Inadequate sampling of length and age composition data from commercial landings and from bycatch of Spanish mackerel from the shrimp fishery.
- Lack of length and age composition sampling from commercial and recreational discards.

Specific comments

Stock structure

Spanish mackerel from US waters within the Gulf of Mexico and to the north of Highway 1 in Monroe County, Florida, which have been designated the "Gulf of Mexico stock", were the subject of the stock assessment. The Data Workshop Report acknowledged that studies of stock structure for Spanish mackerel in the Gulf of Mexico and off the US South Atlantic coast have produced conflicting results. The Report advised that, while early morphometric, meristic, allozyme, and electrophoresis studies and a more recent study of otolith shape and chemistry identify differences between fish from the Gulf of Mexico and those from the South Atlantic coast, a recent mitochondrial and nuclear DNA study did not detect a difference, which suggests at least a small amount of genetic flow between the two regions sufficient to homogenize allele frequencies. Based on results of the earlier studies, and taking into account spawning locations, stock distribution patterns, and catch history, the two groups of fish were recognized as separate management units, with a boundary at US Highway 1 in Monroe County, Florida, which has served as the boundary for data collection from the commercial and recreational fisheries. The evidence supporting the proposed stock structure and, in particular, the boundary separating the two putative stocks is not strong. Further studies to improve understanding of stock composition, *e.g.,* genetic, otolith microchemistry, species composition of parasites, tagging studies, should be initiated.

In the review of data relating to stock structure for Spanish mackerel, the Data Workshop Report makes no mention of the southern boundary of the putative Gulf of Mexico stock, and whether this stock extends into Mexican waters. If such extension is the case, failure to take into account Mexican catches of Spanish mackerel would result in bias in assessment results. The stock assessment that has been undertaken implicitly assumes that the Gulf of Mexico stock of Spanish mackerel is confined to US waters, and thus conclusions from the assessment must be considered conditional on the validity of this assumption.

Biological data

The use of Hoenig's (1983) equation for fish and maximum age to produce an estimate of natural mortality *M* for a fish stock is accepted practice when no data are available from the stock to allow direct estimation of this parameter. Thus, noting also that other methods of estimating *M* from life history data were investigated, its use of Hoenig's (1983) equation to estimate the base value of *M* for Gulf of Mexico Spanish mackerel is endorsed. The Data and Assessment Workshops also correctly recognized that this estimate of *M* was imprecise, and that the results of stock assessment were likely to be sensitive to this uncertainty.

For the reasons noted earlier when discussing the assessment for Gulf of Mexico cobia, use of the Lorenzen (1996) equation to derive age-dependent estimates of natural mortality *M* for Gulf of Mexico Spanish mackerel is not endorsed. In his report to the CIE on the stock assessments conducted for yellowtail flounder and Atlantic herring at Woods Hole in 2012, Francis (2012) advised that prediction of *M*, and, through body weight, its variation with age for an individual species, using Lorenzen's (1996) equation was likely to be highly imprecise, as was evident in the wide scatter about the regression line in Lorenzen's Figure 1. Francis observed that, for about one-third of Lorenzen's data points, predicted and observed *M*s appeared to differ by a factor of more than 2. Furthermore, in the case of both herring and yellowtail, the values of *M* estimated by Lorenzen's equation differed markedly from the values estimated using Hoenig's (1983) equation and had to be scaled substantially for use in the yellowtail flounder and Atlantic herring assessments. Francis (2012) raised the very valid point that, if the estimates produced for a species by Lorenzen's equation provide such unreliable estimates that the mean *M* differs from the estimate calculated using Hoenig's (1983) equation by a factor that differs markedly from 1, can it be considered sufficiently reliable to estimate how *M* varies with age within these species?

There has been no test to assess whether the introduction of the additional complexity associated with age-dependent natural mortality to the model for Gulf of Mexico Spanish mackerel is justified by the resultant improvement in fit that was obtained. It is recommended that a model employing a constant value of *M* is fitted to the Spanish mackerel data. If this model fits just as well as the model that employs an age-dependent *M*, then the simpler model should be used. If the age-dependent model produces a better fit, it would be better to estimate age-dependent *M* within the assessment model rather than assuming that it is of the form predicted by the Lorenzen (1996) equation.

Data on the rate of mortality for discarded hook and line caught Spanish mackerel are limited, and thus the estimates of discard mortality are imprecise. It was pleasing to note that the Assessment Workshop investigated the implications of uncertainty in the estimate of discard mortality by conducting a sensitivity run. Further research is required to produce a more reliable estimate.

Although only the parameter estimates of the von Bertalanffy growth curve fitted to the length at age data using the Diaz *et al.* (2004) model are input to Stock Synthesis to provide the initial values of the growth curve fitted within the assessment model, the growth curve developed for the Data Workshop is of value as a basis of comparison with the growth curve fitted by Stock Synthesis. Fitting the growth curve within Stock Synthesis ensures that the assumptions regarding selectivity are consistent with those employed in other parts of the model and that uncertainty in the estimates of growth is reflected in the estimates of the spawning stock biomass, fishing mortality and benchmarks.

Spanish mackerel exhibit dimorphic growth, yet the Stock Synthesis model considers only pooled data. In future refinement of the model, consideration should be given to modelling both females and males rather than combined sexes, noting that the benefit of this might only be realised if appropriate sex composition data for landings and discards are available for input, and length and age-at-length compositions are sexually disaggregated.

The Data Workshop Report advises that, due to a paucity of age data, percentage maturity was related to size class rather than age. It is not clear whether the data reported in Tables 2.3 and 2.4 represent only fish collected during the spawning season, *i.e.*, when mature fish can be distinguished readily from immature fish on the basis of macroscopic examination of their gonads. It is unclear how the age at 50% maturity for females was estimated, *i.e.,* was this obtained by transforming from length to age using the fitted growth curve. Further details are required. The value of 0.2 y seems surprisingly low for the age at 50% maturity of females. This low value drew comment from the Data Workshop, which suggested that it might have been due to identification of mature fish using macroscopic examination and recommended the use of the age at 50% maturity that was determined for the Atlantic stock of Spanish mackerel, *i.e.*, 0.7 y. Using the relationship between age at maturity and maximum age determined by Froese and Binohlan (2000), a species with an age at maturity of 0.2 y would be expected to have a maximum age of 0.8 y, a value far lower than the 11 years that the Data Workshop employed when estimating *M*. Further research to determine the relationship between percentage mature and age appears to be necessary given this unusually low value and the statement in Section 2.8 of the Data Workshop Report that there is a paucity of age data for Gulf of Mexico Spanish mackerel.

Commercial landings

The decision to extend the time series of landings data as far back in time as possible was endorsed, although it is noted that (1) the data in Table 3.2 of the Data Workshop Report were very sparse until 1927, and (2) the reliability of commercial data improved substantially in 1950. Note that it would be useful to state in the heading of Table 3.2 whether the gaps in data prior to 1927 represent missing years, or, as reported in Table 3.4, represent zero landings. As an alternative to using data extending back to 1887, it might be interesting to compare the results obtained from the model by using a shorter time series ranging from 1927 to 2011, noting that the imprecision associated with imputing the missing landings between 1887 and 1926 should also be considered.

The decision made by the Data Workshop to combine landings from commercial fishing gears other than gillnets and handlines was not explained. Was it to reduce the number of time series of landings considered in Stock Synthesis, and thereby reduce complexity, or was the decision made in recognition of a lack of data to characterize the length composition of each of the miscellaneous gears? A decision made because of the latter reason would indicate an inadequacy of the data collection programs, which might need to be addressed.

Until 1996, the annual landings of the combined commercial gears, other than gillnets and handlines, were typically of a greater magnitude than the landings made by handlines, and subsequently were of similar magnitude. As recommended by the Data Workshop, the Assessment Workshop apportioned these combined landings of the miscellaneous commercial gears to the landings of the two primary gears in proportion to the annual landings of those last two gears. The length composition of the resultant time series of landings thus reflect a weighted combination of the length compositions of the catches from the different fishing gears, each of which would have reflected the selectivity curve of that gear. Length composition data collected from the landings taken using gillnets or those taken using handlines will therefore fail to reflect the length compositions of the mixtures of landings of those primary gears and the contribution from the landings of the miscellaneous gears, particularly in the case of the length composition data for the handline landings.

Comment is made in Section 3.3.5 of the Data Workshop Report that there was a precipitous decrease in landings in 1977 and subsequent years following cold weather in Florida in 1976-77. This environmental event was not explored by the Assessment Workshop, but it might be interesting to consider whether the cold weather caused increased mortality or reduced growth, and whether this could explain the reduced landings that followed the 1977 event.

The Data Workshop is commended for its collation of the commercial landings data from the various sources and development of a time series of commercial landings suitable for use in the stock assessment process for the Gulf of Mexico stock of Spanish mackerel. It would be useful to assess and report the imprecision of the annual estimates.

Although the Data Workshop Report advised that the decision was made that discarded fish, which were designated as “kept”, should be removed from the amount of discards and added to landings, it is unclear whether this was done when preparing the landings and discard data for the Assessment Workshop.

Discards recorded for the commercial fisheries are highly uncertain due to low reporting rates and are likely to represent minimum values. Programs to collect discard data from commercial fishers need to be reviewed to identify ways in which more reliable discard data might be obtained.

The Bayesian model, which assumed that counts within cells had a negative binomial distribution, appeared an appropriate approach to estimating the bycatch of Spanish mackerel by the shrimp fishery. The Data Workshop advised, however, that, as a consequence of low encounter rate of Spanish mackerel by the shrimp fishery and irregular observer coverage, estimates of bycatch of Spanish mackerel are imprecise, although the mean is likely to be of the appropriate scale.

The Data Workshop Report advised that “sample sizes for developing length compositions were inadequate for a considerable number of year and gear strata”. Sampling to determine the age compositions of commercial landings has also been sparse, particularly for gillnet landings in recent years. There appear to be no data that could be used to characterize the length or age compositions of discards from the commercial fisheries. Data collection programs should be reviewed to identify how they could be improved to collect representative samples of length and age compositions from the landings and discards of the commercial fisheries.

Recreational landings

As with the commercial landings data, the Data Workshop is commended for its collation of the recreational landings of Gulf of Mexico Spanish mackerel from the various data sources, and, in particular, the extension of this time series of data back to 1955.

The Assessment Workshop reported that the estimates of discards of Spanish mackerel from the recreational fishery were highly uncertain, due to low intercept rates and the changes in quality control and assurance that had occurred between 1981 and 2011.

Age samples for the recreational fishery were collected by the Southeast Region Headboat Survey (SRHS), as lengths but not ages are typically collected within the MRFSS. No samples were available to characterize the length and age compositions of discards of Spanish mackerel by recreational fishers. Consideration should be given to developing a program to collect representative length and age data from Spanish mackerel that are discarded by the recreational fishery.

Survey indices

The recommendation reported in the Data Workshop Report that the fishery-independent SEAMAP survey and the fishery-dependent MRFSS, and FL trip ticket handline/trolling indices, are appropriate for use in the assessment, and that other putative indices should not be used, appears sound. Both the SEAMAP and MRFSS surveys used a delta lognormal model to standardize the data and thereby determine annual indices of abundance. The trip ticket data were standardized using a general linear model with forward stepwise selection.

In Section 5.4.4.6 of the Data Workshop Report, the Working Group advised that the index of abundance based on data from headboats was adequate for use in the assessment, yet the report card for the index advises that, because of the small proportion of observations that reported catches of Spanish mackerel, the Working Group did not endorse the use of the index in the assessment. Table 5.4.4.1 in the Data Workshop Report incorrectly divides total trips by total positive trips and reports the result, 38.89, as the overall percentage of positive trips instead of 2.6%. The incorrect value is then taken from the table and reported as 38.89% in Section 5.4.4.2 of the Data Workshop Report. The overall summary in section 5.1 correctly advises that the headboat index was not recommended for use. Accordingly, the Assessment Workshop did not include this as a survey to be used by Stock Synthesis.

ToR 11. Evaluate the quality and applicability of methods used to assess the stock.

Conclusions

Stock Synthesis 3, the software within which the model for the Gulf of Mexico stock of Spanish mackerel was developed, has gained international recognition for its quality and the applicability of the methods it uses to assess the condition of fish stocks. The model for Spanish mackerel was of an appropriate structure given the data that were available. Values predicted by the model for Spanish mackerel, including those of benchmarks, were imprecise, however, due to the nature of the input data. Further imprecision of model outputs due to alternative values of key parameters, such as natural mortality and steepness of the stock-recruitment relationship, was explored. Recognising the types of data that were available for input and the uncertainty of model outputs that arose as a consequence of the nature of those input data, the Stock Synthesis base model for Spanish mackerel is of a quality consistent with that which would be considered "best practice", and is able to provide a valuable assessment of the likely condition of the stock in 2011, and, when projected, the likely trajectory of yields and stock condition over the next five to six years.

Strengths

- The decision to use Stock Synthesis 3 as the modelling framework and to complement this with the Fishery Simulation Graphics User Interface (Lee *et al*., 2012).
- The structure of the model developed within the Stock Synthesis framework was appropriate given the data that were available.
- The enhancement of Stock Synthesis to allow modelling of a fishery for which the only source of mortality is that associated with discarding of bycatch.
- Use of super periods when data are too imprecise to fit individual values but the median value is considered to be informative.
- The assessment of the uncertainty of parameter estimates was thorough.
- Selectivity runs explored key uncertainties and demonstrated appropriateness of conclusions regarding the current condition of the stock.
- Benchmarks were appropriately calculated.
- Projections were undertaken using two states of nature.

Weaknesses

- Subjective decision to set effective sample size to actual sample size capped at a maximum of 100 rather than to use iterative reweighting, such as proposed by Francis (2011).
- Lack of information in abundance indices, and shortness of history of length and age-at-length data.
- Lack of length and age composition data to provide information on the length and age compositions of discards and the shape of the retention curves.
- The assumption that natural mortality is age-dependent and has a form that is proportional to the values predicted by the Lorenzen (1996) has not been tested against the simpler assumption of constant natural mortality over age.
- Imprecision in the estimate of steepness of the stock-recruitment relationship.
- Lack of exploration of uncertainty associated with the time series of commercial and recreational landings.

The assessment was undertaken using Stock Synthesis 3, a fully integrated model that allowed use of all available data for Spanish mackerel in the Gulf of Mexico, including life history data, removals, discards, length compositions of catches, conditional age-at length compositions, and survey indices. Other software packages, which were used in the assessment of the Gulf of Mexico Spanish mackerel stock, were r4SS, which produces graphic displays and explores output from Stock Synthesis, and the "Fishery Simulation" Graphics User Interface (GUI) software (Lee *et al*., 2012), which adds bootstrapping analysis support to Stock Synthesis. Stock Synthesis, supported by these software packages, provides a very flexible assessment framework that produces estimates of key population parameters and their uncertainty. The software allowed exploration of the sensitivity of parameters, stock status indicators, and reference points to changes in the structure of the Spanish mackerel model and its assumptions, and to the exclusion of various survey indices when fitting. It also allowed investigation of yield per recruit, spawner per recruit, and stock-recruitment relationships for Spanish mackerel, and produced estimates of reference points to be used when determining stock status. The Stock Synthesis model was also employed to project the effect of different levels of fishing mortality on future catches and condition of the Gulf of Mexico Spanish mackerel stock. Through bootstrapping, Stock Synthesis was used to develop probability distributions for various variables of interest.

The Assessment Workshop Report advised that, apart from the FWC trip ticket vertical line index, which showed a slight increase in abundance after 2003, predicted values of the abundance indices, which exhibited considerable imprecision, were relatively constant over the periods for which abundance indices were available. As noted by the Assessment Workshop, this implies that the survey indices carry little information regarding trends in abundance. The Assessment Workshop also noted that length and conditional length-at-age data cover only a limited recent period, and thus provide limited information on recruitment to inform the model.

Concern that the estimate of steepness produced when fitting the initial model, *i.e.*, 0.52, was too low, led the Assessment Panel to profile log-likelihood over a range of values of steepness (Fig. 3.31, Assessment Workshop Report), thereby to assess whether the data were sufficiently informative to allow reliable estimation of this parameter. After examining the results of this and other sensitivity runs, retrospective analyses, profiling, and bootstrap runs, the Assessment Panel concluded that a value of 0.8 for steepness “was more reasonable for this species than that estimated by the model (0.52)” (see further comment regarding this decision below), and adopted this configuration (Run 3) as the base model for the assessment. That is, Run 3 was recommended by the Assessment Panel for final projections and status determinations.

The use within Stock Synthesis of super periods when fitting discards of Spanish mackerel from the commercial line gear fishery, the recreational fishery, and the shrimp fishery, is very appropriate given the high uncertainty associated with the estimates of the annual discards for these three fisheries. By fitting estimates of discards to the average value of discards over these super periods, the model “accepts” the overall level but “ignores” inter-annual variability within the discard time series.

The assumption that was made in the assessment that age data were conditional on length is very appropriate. If it had been assumed that the length and age composition data were independent, the fact that some fish were included in both the length and age composition data would introduce bias. Such potential bias is removed by considering ages to be conditional on length.

The decision that, because of a lack of strong evidence that selectivity was dome-shaped and the fact that little improvement in fit was obtained when using such a selectivity pattern, selectivity functions for the commercial line gears and recreational fisheries would be constrained to those with an asymptotic pattern is endorsed. It was good to note that some exploration had been undertaken before coming to this conclusion, but it would have been useful if the results of that exploration had been presented in the Assessment Workshop Report. The representation of the retention curves using two time blocks, *i.e.* the period before 1993 and the period from 1993 onward, to reflect the change in size limit in 1993, is appropriate.

It would have been appropriate to explore whether the improvement in likelihood of the fitted model justified the additional complexity of considering mortality to be age dependent rather than constant. If not justified, the simpler model would be preferred. If use of an age-dependent model was justified, it would be better to estimate the values of the age-dependent mortalities directly, rather than assuming that the relationship has a form that is a scaled version of the values of mortality at age calculated using Lorenzen's (1996) equation.

The use of a maximum effective sample size of 100 fish is arbitrary, however, it is noted that Sensitivity Run 12 explored the effect of reweighting using the MacAllister and Ianelli (1997) approach. It is recommended that, in future analyses, consideration should be given to the methods described by Francis (2011), such that, for example, effective sample sizes for length compositions are calculated using iterative reweighting based on mean length, and possibly reflecting the relative magnitudes of initial sample sizes.

No length or age composition data were available to characterize the discards from the commercial or recreational catches, thus little information was available to estimate the parameters of the logistic retention curves for these fisheries.

The use of a Beverton and Holt stock-recruitment curve is endorsed, but the choice of the value of 0.7 as the value of the standard deviation in recruitment appears arbitrary. The Assessment Workshop Report advised that the profile of likelihoods over a range of values "did not indicate disparity" with the value chosen (Fig. 3.33). It might be pertinent to note, however, that both Smith and Punt (1998) and Maunder and Deriso set $\sigma^2_{\log_e R} = 0.6$. Beddington and Cooke (1983) are cited as reporting from a meta-analysis over many fish species that recruitment is typically log-normally distributed with the average of $\sigma^2_{\log_e R}$ being around 0.6. Mertz and Myers (1996) are reported to have conducted a further meta-analysis and again found that the average value of $\sigma^2_{\log_e R}$ was around 0.6. Interestingly, the likelihood profile (Fig 3.33) suggests that 0.6 might be slightly more appropriate than 0.7.

As advised in the Assessment Workshop Report, Stock Synthesis effectively treats landings as being known without error and thus fits them precisely. Imprecision associated with the early values within the time series of commercial or recreational landings is thus not assessed unless explored through sensitivity runs using alternative scenarios of landings data. It is not apparent from the Assessment Workshop Report that such sensitivity runs were made and thus the implications of the uncertainty associated with the landings data have not been assessed.

In describing Fig. 3.35, it is unclear whether the 14 of the 1000 bootstrap runs, which produced "large convergence values and illogical estimates of virgin biomass" were not simply the results of poor choices of initial values for the parameters used in Stock Synthesis, given that the jitter analysis produced four out of 100 results that failed to converge to the expected values.

The vertical scale used in the profile of change in log-likelihood over the range of values of steepness (Fig. 3.31, Assessment Workshop Report) compresses the range of values of log-likelihood change for values of steepness ranging from (say) 0.4 to 0.9, which is the region of interest. A maximum value on the y-axis of (say) 100, would have more clearly revealed the trend in log-likelihood change.

The conclusion by the Assessment Workshop that the estimate of steepness is imprecise is valid, however, although the range of values that, given the model structure and data, might be considered to fall within a 95% confidence region would probably extend from about 0.4 to about 0.8. The basis for the decision by the Assessment Panel that a value of steepness of 0.8 is "more reasonable" than the estimated value of 0.52 for the Gulf of Mexico stock of Spanish mackerel is not stated. In this context, it is possibly pertinent to note that Francis (2012) has suggested that, when the steepness of the stock-recruitment relationship is imprecise or cannot be estimated reliably, he considers it better to fix the value of steepness at a value, such as 0.75, *i.e.*, the default value recommended in Francis (1993), and which is frequently used in Australia and New Zealand, or the average of published values for the same or similar species. Francis (2012) advises that the uncertainty associated with this parameter should then be explored using sensitivity runs with lower and higher values of steepness. The value of steepness selected by the Assessment Workshop, *i.e.*, 0.8, is of similar magnitude to the value suggested by Francis (2012), *i.e.*, 0.75. Thus, the decision by the Workshop to use a model with a structure similar to that of the original base model but with a fixed value of steepness of 0.8, *i.e.*, the model of Run 3, as the new base model for the Spanish mackerel stock, and to explore the uncertainty associated with this steepness using sensitivity runs with alternative values of steepness, is consistent with best practice, and is therefore endorsed.

The use of the base model, and of a model with similar structure but with steepness fixed at 0.9, as alternative states of nature is endorsed. Given the results of the sensitivity runs, however, it might also have been useful to include a low natural mortality version of the base model as a third state of nature.

ToR 12. Recommend appropriate estimates of stock abundance, biomass, and exploitation.

Conclusions

Estimates of stock abundance, biomass, and exploitation are produced when the Stock Synthesis model is fitted. The estimates of total biomass and annual exploitation in 2011, which were estimated when the base model for the Gulf of Mexico stock of Spanish mackerel was fitted, were 28,367 mt and 0.1197, respectively.

Strengths

Stock Synthesis 3 calculates time series of abundance, total biomass, and annual exploitation.

Stock abundance:

The report file that is produced by Stock Synthesis, report.sso, contains a time series section, in which the time series of abundance, recruitment and catch for each of the areas are reported. Output quantities include summary biomass and summary numbers for each gender and growth pattern. The Assessment Workshop Report for the Gulf of Mexico Spanish Mackerel stock has not reported these abundance estimates, but they will be available in the output file for the base model, *i.e.*, Run 3.

Biomass:

Stock Synthesis produces an estimate of total annual biomass (Table 3.5, Fig. 3.41). The estimate (for the base model, *i.e.,* Run 3) of total biomass for 2011 was 28,367 mt.

Exploitation:

Stock synthesis calculates the value of annual exploitation rate as the ratio of the weight of the total catch (including discards) to the total biomass (Section 3.26, Assessment Workshop Report; Table 3.6, Fig. 3.42). The calculated value of the annual exploitation rate is used as a proxy for the annual value of fishing mortality, *F*. The estimate (for the base model, *i.e*., Run 3) of the annual exploitation rate for 2011 was 0.1197.

ToR 13. Evaluate the methods used to estimate population benchmarks and management parameters. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.

Conclusions

Stock Synthesis calculates a range of population benchmarks and management parameters. Benchmarks calculated for Spanish mackerel were MFMT = $F_{30\%\mathrm{SPR}}$ and MSST = (1 – *M*) $\mathrm{SSB}_{30\%\mathrm{SPR}}$. The estimates of F_{current} and $\mathrm{SSB}_{\mathrm{current}}$, which were calculated for 2011 using the base model, were 0.14 and 19,645 mt, respectively. The ratios $F_{\mathrm{current}}/\mathrm{MFMT}$ and $\mathrm{SSB}_{\mathrm{current}}/\mathrm{MSST}$, which were calculated using the base model, were 0.38 and 3.06, respectively. These results, which were consistent with those produced by all but one (the model with natural mortality set to 0.27 y^{-1}) of the models used in the various sensitivity runs, imply that, in 2011, the Gulf of Mexico stock of Spanish mackerel was not experiencing overfishing and was not overfished.

Strengths

Stock Synthesis possesses well-tested procedures to calculate and output a range of population benchmarks and management parameters.

Weaknesses

Inconsistencies in the values recorded in one of the columns in Table 3.8 made it difficult to assess, with full confidence, whether or not the stock was experiencing overfishing.

Summary

The methods used by Stock Synthesis to estimate population benchmarks and management parameters are sound. Stock Synthesis is able to produce estimates for indicator variables and reference points based on maximum sustainable yield (MSY), spawning potential ratio (SPR), and spawning stock biomass (SSB), and taking the stock-recruitment relationship into account. SPR is calculated as the equilibrium spawning biomass per recruit that would result from a given year's pattern and the levels of F's and selectivities for that year. For MSY-based reference points, Stock Synthesis searches for a fishing mortality that would maximise the equilibrium yield. For SPR-based reference points, the computer program searches for an F that would produce the specified level of SPR. For spawning biomass-based reference points, the software searches for an F that would produce the specified level of spawning biomass relative to the unfished value.

The management benchmarks, *i.e.*, the Maximum Fishing Mortality Threshold (MFMT) and Minimum Stock Size Threshold (MSST), which were proposed for the fishery by the Assessment Workshop, are appropriate for use in determining the status of the Gulf of Mexico stock of Spanish mackerel. These two benchmarks were

$$\text{MFMT} = F_{\text{MSY}} \quad \text{and} \quad \text{MSST} = (1 - M)\ \text{SSB}_{\text{MSY}},$$

where F_{MSY} is the fishing mortality that produces the maximum sustainable yield MSY, M is the point estimate of natural mortality for fully recruited ages calculated using Hoenig's (1983) equation, *i.e.* 0.38 y^{-1}, and SSB_{MSY} is the equilibrium spawning stock biomass that produces MSY. The Assessment Workshop Report advises that proxies were used when calculating the above benchmarks, where these proxies were based on a spawning potential ratio (SPR) of 30%. Thus, the proxy that was used for F_{MSY} was the fishing mortality, $F_{30\%\text{SPR}}$, which produces a spawning stock biomass per recruit that is 30% of the spawning stock biomass per recruit produced when the stock is not fished, *i.e.* an SPR of 30%. The proxy that was used for SSB_{MSY} was the corresponding value of equilibrium spawning stock biomass, *i.e.* the spawning stock biomass $\text{SSB}_{30\%\text{SPR}}$ that is produced with a fishing mortality of $F_{30\%\text{SPR}}$.

It is surprising to note that, although Stock Synthesis was able to estimate MSY-based rather than SPR-based reference points, the Assessment Panel chose to use the proxies $F_{30\%\text{SPR}}$ and $\text{SSB}_{30\%\text{SPR}}$ rather than F_{MSY} and SSB_{MSY}. The latter two benchmarks are possibly more appropriate.

For the Gulf of Mexico stock of Spanish mackerel, the benchmarks that were used in determining stock status by the Assessment Workshop were

$$\text{MFMT} = F_{30\%\text{SPR}} \quad \text{and} \quad \text{MSST} = (1 - M)\ \text{SSB}_{30\%\text{SPR}},$$

where it was concluded that overfishing was occurring if $F_{\text{current}} > \text{MFMT}$, *i.e.*, $F_{\text{current}}/\text{MFMT} > 1$, and the stock was considered to be overfished if $\text{SSB}_{\text{current}} < \text{MSST}$, *i.e.*, $\text{SSB}_{\text{current}}/\text{MSST} < 1$. F_{current} was calculated as the geometric mean of the estimates of the three most recent annual fishing mortalities, *i.e.,* the fishing mortalities for 2009-2011, where annual fishing mortality was estimated by its proxy, exploitation rate, calculated as the ratio of the total catch (including discards) to estimated total biomass. $\text{SSB}_{\text{current}}$ was the estimate of spawning stock biomass for 2011.

Note that the specification of the reference points in Section 3.1.9 of the Assessment Workshop Report could be improved, *e.g.* overfished is currently defined as the value of the ratio of $\text{SSB}_{\text{current}}$ to MSST rather than a logical expression.

Table 3.8 of the Assessment Workshop Report, which is reproduced below, contains the values of the current (2011) fishing mortality and spawning stock biomass of the Gulf of Mexico stock of Spanish mackerel, and purports to contain the values of the MFMT and MSST benchmarks, and the results of stock determination for each of the models that were explored in the assessment. According to the caption for this table in the Assessment Workshop Report, F_{ref} represents $F_{30\%\mathrm{SPR}}$, and thus, as MFMT has been set to $F_{30\%\mathrm{SPR}}$, the values of MFMT should be equal to those of F_{ref}. As is evident in Table 3.8, this is clearly not the case. There are inconsistencies between the values of F_{ref} and MFMT for all but three of the 17 runs presented in the Table, Quite frequently, however, the values of F_{ref} and the ratio of F_{current} to MFMT in the rows of this Table are equal. The caption to Figure 3.9 advises that, for this figure, the value of F_{ref} represents the ratio of F_{current} to MFMT, and it appears likely that this inconsistency between definitions of F_{ref} has led to the inconsistent values presented in Table 3.8. The fact that there is such inconsistency makes it difficult to accept the accuracy of the estimates of the ratio of F_{current} to MFMT for any of the runs. Accordingly, while it is not possible from the reported data to assess with complete confidence whether or not the stock is experiencing overfishing, if the values in the column headed "F/MFMT" are correct, then $F_{\mathrm{current}}/\mathrm{MFMT} = 0.38$. From this, and noting the values for this ratio for other selectivity runs, it is very likely that the Gulf of Mexico stock of Spanish mackerel is not currently being subjected to overfishing.

Quoted from Assessment Workshop Report: "Table 3.8. Reference points and benchmarks from sensitivity runs for Gulf of Mexico Spanish mackerel from SS. Benchmarks are reported for SPR 30%. Current refers to geometric mean of 2009-2011 values. MSST is (1-M)*SSBref with M = 0.38, or M=0.27, or M=0.49 representing the M value from the Hoenig maximum age mortality estimator for fully recruited ages from the SEDAR DW corresponding to the Base Model M or the M_LO or M-HI scenario. Ref refers to reference metric, either F30% SPR or SSB 30% SPR. Fratio is Fcurrent / Fref. SSBratio is SSBcurrent / MSST. Spawning biomass units are weight in mtons, and yield units are mtons whole weight".

Name	Fcurrent	SSBcurrent	Yref	Fref	SSBref	MFMT	MSST	F/MFMT	SSB/MSST
Run 1 Configuration	0.19	11,195	3,563	0.37	6,626	0.37	4,108	0.51	2.73
Run 1 Configuration, Steepness=0.9	0.14	18998	3090	0.39	10701	0.35	6634	0.39	2.86
Run 1 Configuration, Steepness=0.8	0.14	19,645	3,053	0.39	10,339	0.36	6,410	0.38	3.06
Run 1 Configuration, Steepness=0.7	0.14	18,235	3,056	0.41	10,264	0.35	6,363	0.41	2.87
Run 3 Configuration, M HI	0.1	23,551	3,682	0.2	8,746	0.5	4,461	0.2	5.28
Run 3 Configuration, M LO	0.2	13,150	4,040	0.83	18,283	0.24	13,347	0.83	**0.99**
Run 3 Configuration, M REF Age 3	0.15	18,140	3,138	0.47	11,862	0.32	7,354	0.47	2.47
Run 3 Configuration, Discard Mortality	0.14	18,995	3,029	0.41	10,730	0.35	6,653	0.41	2.86
Run 3 Configuration, NO MRFSS	0.14	19,886	3,054	0.39	10,637	0.35	6,595	0.39	3.02
Run 3 Configuration, NO FWC	0.12	25,700	2,821	0.34	11,132	0.34	6,902	0.34	3.72
Run 3 Configuration, NO SEAMAP Survey	0.13	20,364	3,053	0.38	10,715	0.35	6,643	0.38	3.07
Run 1 Configuration, SS Reweighting	0.19	11,050	3,743	0.37	7,011	0.37	4,347	0.5	2.54
Run 3 Configuration, RETROSPECITVE_2010	0.15	18,383	3,163	0.43	10,882	0.35	6,747	0.43	2.72
Run 3 Configuration, RETROSPECTIVE_2009	0.16	17,503	2,991	0.46	11,022	0.34	6,834	0.46	2.56
Run 3 Configuration, RETROSPECTIVE_2008	0.15	18,121	2,968	0.44	11,182	0.35	6,933	0.44	2.61
Run 3 Configuration, RETROSPECTIVE_2007	0.15	16,832	3,072	0.46	11,362	0.33	7,044	0.46	2.39
Run 3 Configuration, RETROSPECTIVE_2006	0.16	19,528	3,040	0.48	10,986	0.34	6,811	0.48	2.87

The point estimates of the ratio of $\mathrm{SSB}_{\mathrm{current}}/\mathrm{MSST}$ exceed 1 in all but one case of Table 3.8 of the Assessment Workshop Report, *i.e.*, that for the run in which *M* was set at the lower value, MLO = 0.27 y^{-1}, when this ratio became 0.99, *i.e.*, the SSB was only just below MSST. Apart from this run, the results of the model runs that were undertaken indicate that that it is highly likely that the stock of Spanish mackerel is currently not overfished.

The value of F_{current} for the model with steepness set to 0.8 is reported as 0.14 in Table 3.8 and 0.13 in Table 3.9 of the Assessment Workshop Report. The ratio of F_{current} to MFMT is reported in Tables 3.8 and 3.9 as 0.38 and, 0.50, respectively for this model, and, for the model with steepness of 0.9, as 0.39 and 0.52, respectively. The values of $\text{SSB}_{\text{current}}$ reported in Table 3.8 for the models with steepness values of 0.8 and 0.9 are transposed in Table 3.9. The values of the ratio of $\text{SSB}_{\text{current}}/\text{MSST}$ in Table 3.9 do not match the values reported in Table 3.8 for either model. These inconsistencies should be resolved.

ToR 14. Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.

Conclusions

Stock Synthesis provides a well-tested procedure to project the model through a range of future years, using a fishing rate based on MSY, SPR, a specified target biomass, or a multiple of the recent average fishing rate and producing estimates of yield and key management parameters, thereby allowing assessment of future stock condition. The methods used, which are recognised as being of high quality, are designed to produce the estimates of future population status that are needed by managers. If the current fishing rate is maintained over the next 10 years, the projections produced for the base model for the Gulf of Mexico Spanish mackerel stock suggest that there will be little change in spawning stock biomass. If, however, fishing mortality is increased to the level that is estimated as required to produce OY, or further increased to that which would produce a spawning potential ratio of 30%, the spawning stock biomass would be expected to be reduced by approximately 20%. The condition of the stock would be expected to continue to be classified as “not overfished, with overfishing not occurring”.

Strengths

Projections are undertaken using the well-tested procedures provided within Stock Synthesis.

Summary

Stock Synthesis includes a well-tested procedure to project the future stock status that would result when using a fishing rate based on MSY, SPR, a specified target biomass, or a multiple of the recent average fishing rate. Use of this procedure ensures consistency of model predictions with assumptions and parameter estimates used in fitting the model and the age structure predicted as the current state of the stock from which the projection commences. It is thus highly applicable for use with the Gulf of Mexico stock of Spanish mackerel.

For the Gulf of Mexico stock of Spanish mackerel, deterministic projections were run by the Assessment Panel for the models with steepness of 0.8 and 0.9 and using fishing rates set to MFMT (*i.e.*, the proxy $F_{30\%\text{SPR}}$ for F_{MSY}), F_{OY} (*i.e.*, 75% of $F_{30\%\text{SPR}}$), and F_{current}. Using the bootstrapping facility provided by the Fishery Simulation GUI software, stochastic projections were also run for the two models with the fishing rate set to MFMT (the Assessment Workshop report only presents the results for the model with steepness set to 0.8).

The final year of the time series of data used in the assessment for the Gulf of Mexico stock of Spanish mackerel was 2011. In order to carry out projections for 20 years from 2013 (only results from 2013 to 2022 being reported), the 2012 landings "were characterized as the landings [of the different fisheries] from the most recent three years (2009-2011)" (Assessment Workshop Report). Stock Synthesis was used to estimate the fishing mortality for 2012 required to achieve these landings, and used the 2012 estimate of SSB to calculate an estimate of age 0 recruitment from the fitted stock-recruitment relationship.

If the current fishing rate is maintained over the next 10 years, the projections produced for the models with steepness set to 0.8 and 0.9 suggest that there will be little change in spawning stock biomass. If, however, fishing mortality is increased to the level that is estimated as required to produce OY, or further increased to that which would produce a spawning potential ratio of 30%, the spawning stock biomass would be expected to be reduced by approximately 20 or 30%, respectively.

ToR 15. Evaluate the quality and applicability 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**
- **If there are significant changes to the base model, or to the choice of alternate states of nature, then provide a probability distribution function for the base model, or a combination of models that represent alternate states of nature, presented for review.**
- **Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments**
- **Provide justification for the weightings used in producing the combinations of models**

Conclusions

The methods within Stock Synthesis that may be used to explore uncertainty include calculation of estimates of asymptotic standard errors, calculation of likelihood profiles, MCMC analyses, and bootstrapping. These tools are complemented by auxiliary software that allows production of diagnostic plots, which also assist in communicating the uncertainty of estimates. The software encourages exploration of alternative model structures and sensitivity to alternative values of parameters of functional forms. The model that was developed for the Gulf of Mexico stock of Spanish mackerel employed an appropriate set of these methods. As a result of the exploration of the uncertainty of the estimate of steepness, the base model was modified by fixing steepness to 0.8. Probability distributions were produced for a set of key parameters using both the original and new base models. As the iterative approach required to calculate *P** cannot be implemented in Stock Synthesis, Stock Synthesis "calculates the expected time series of probabilities that the F resulting from a specified harvest policy would exceed a specified level" (Methot and Wetzel, 2012).

Strengths

- Stock Synthesis provides an extensive suite of methods that may be used to explore uncertainty.
- Bootstrapping was used to produce probability distributions

Summary

Stock Synthesis provides a number of methods that may be used to characterize the uncertainty associated with the estimates of parameters, benchmark estimates, and predicted values of parameters. These are supplemented by the bootstrapping tools provided by the Fishery Simulation GUI. Together, the software is well suited for use in exploring the uncertainty associated with the models that were fitted to the Gulf Of Mexico Spanish mackerel stock. Thus, for each run of the Stock Synthesis model for the Gulf of Mexico Spanish mackerel, asymptotic standard errors were calculated for each of the parameters that were estimated (see Table 3.1, Assessment Workshop Report, for parameter estimates and estimates of asymptotic standard errors for the base model, with M = 0.38 y^{-1} and steepness = 0.8). These estimates of asymptotic standard errors may be considered to represent minimum values for the uncertainty of the estimated parameters. The uncertainty of selected parameter estimates for the Gulf of Mexico Spanish mackerel stock was also characterized using the results from bootstrapping.

The initial run (Run 1) was carried out using the model structure that had been proposed for the Gulf of Mexico stock of Spanish mackerel and estimating the steepness parameter of the Beverton and Holt stock-recruitment relationship. This demonstrated that, given the data that were available, the steepness of the stock recruitment relationship was estimated very imprecisely. A number of sensitivity runs of Stock Synthesis were then run to explore the effect of varying the configuration or methods employed in the analysis.

As is typical in stock assessment, exploratory runs for the Gulf of Mexico Spanish mackerel stock were first employed to determine a base model for the assessment, *i.e.*, a model that is considered the most likely of the alternative model configurations that have been proposed. The decision was made at the Assessment Workshop to reject Run 1 and use Run 3 as the base model. As noted above, a justification for this decision, *i.e.*, to use the initial model structure, *i.e.*, that for Run 1, and to fix the value of steepness at 0.8, was not reported in the Assessment Workshop Report other than to state that the Assessment Workshop found the low estimate of steepness produced when fitting the model in Run 1 to be unacceptable. Probability distributions of the key parameters estimated for the initial model, Run 1, and the new base model, Run 3, were produced and plotted (Figs 3.34 and 3.35 of the Assessment Workshop Report).

The level to which the initial spawning stock biomass had been depleted by 2011 was far less for Run 1, *i.e.*, 0.16 SSB_{B0} than for Run 3, *i.e.*, 0.51 SSB_{B0} (Table 3.7, Assessment Workshop Report). A similar level of depletion, *i.e.*, 0.18 SSB_{B0} as that of Run 1 was estimated to have resulted when the value of natural mortality used in the Run 3 configuration was lowered to 0.27 y^{-1}. When Run 1 was re-fitted, estimating steepness (with a resulting value of 0.53) and iteratively adjusting the weights of the survey indices and the length and age compositions to match the estimated variances of the input data with those of the fitted model, the level of depletion was again low, *i.e.*, 0.16 SSB_{B0}. The level of depletion of spawning stock biomass appears sensitive to reduced values of steepness and/or natural mortality. Given the estimated level of depletion of spawning stock biomass for these runs, it is interesting to note that SPR had been reduced in these three model configurations to only 0.51, 0.41, and 0.53, respectively (Table 3.7, Assessment Workshop Report). Again, these results suggest that, when MSY-based reference points are available, these should be used in preference to SPR-based proxies.

While the Assessment Workshop Report provided a comparison of the key parameters, benchmarks, and projections for the base model that was adopted at the workshop, *i.e.*, Run 3, with steepness of 0.8, and an alternative model, which had an identical configuration but used a steepness of 0.9, the relative probabilities of the two models was not assessed. The base model was subjected to a bootstrapping analysis, however, and distributions of the resulting estimates of the benchmark estimates are provided in Figures 3.48 and 3.49 of the Assessment Workshop Report, while distributions of projected yields for 2013-2022 are plotted in Fig. 3.53.

The caption of Table 3.9 advises that the table provides results of the required SFA and MSRA evaluations using a SPR 30% reference point for "4 states of nature of steepness at 3 levels of natural mortality". The table, however, only presents results for models representing two values of steepness for one value of natural mortality.

While the iterative approach required to calculate P^* cannot be implemented in Stock Synthesis, a complementary approach has been developed to produce estimates of the probability that F, the fishing rate based on MSY, SPR, a specified target biomass, or a multiple of the recent average fishing rate that is employed in the projection, exceeds the OFL (Methot and Wetzel, 2012). These authors advise that, whereas the P^* approach calculates the future stream of annual catches that would have a specified annual probability of $F > \text{OFL}$, Stock Synthesis "calculates the expected time series of probabilities that the F resulting from a specified harvest policy would exceed a specified level".

ToR 16. If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.

The Review was undertaken as a desktop review, rather than in a Workshop setting. Accordingly, it was not possible for the recommendations made in review reports to be acted upon, nor to ensure that the results were incorporated accurately in the resultant Stock Assessment Report.

ToR 17. Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.

The SEDAR Process has ensured that all aspects of the assessment process for the Gulf of Mexico stock of Spanish mackerel, from collation of data through to model development, exploration, and production of management advice, have been documented in detail, including the underlying reasons for the decisions that were made concerning data to be used and model structure to be employed. The structure imposed on the Data and Assessment Workshops by their Terms of Reference has assisted by providing a logical framework for the process, and thereby ensuring that key aspects of the assessment were not overlooked. For the reviewer, the documentation of the Spanish mackerel assessment, which was produced through the SEDAR process, proved invaluable in gaining an understanding of the details of the assessment and assisted in identifying opportunities for improvement and in detecting errors or inadequacies.

The Terms of Reference for the Assessment Process, which are presented below, are now examined and comment is made on the degree to which these were addressed.

1. Review and provide justification for any changes in data following the data workshop and any analyses suggested by the data workshop. Summarize data as used in each assessment model.

   Accomplished.

2. Recommend a model configuration which is deemed most reliable for providing management advice using available compatible data. Document all input data, assumptions, and equations.

   Accomplished.

3. Incorporate known applicable environmental covariates into the selected model, and provide justification for why any of those covariates cannot be included at the time of the assessment.

   No environmental covariates were identified by either the Data or Assessment Workshops.

4. Provide estimates of stock population parameters.
   - Include fishing mortality, abundance, biomass, selectivity, stock-recruitment relationship, and other parameters as appropriate given data availability and modeling approaches
   - Include appropriate and representative measures of precision for parameter estimates

   Accomplished.

5. Characterize uncertainty in the assessment and estimated values.
   - Considering components such as input data, modeling approach, and model configuration
   - Provide appropriate measures of model performance, reliability, and ‘goodness of fit’

   Accomplished.

6. Provide yield-per-recruit, spawner-per-recruit, and stock-recruitment evaluations.

   Accomplished.

7. Provide estimates of stock status relative to management criteria consistent with applicable FMPs, proposed FMPs and Amendments, other ongoing or proposed management programs, and National Standards for each model run presented for review.

   Accomplished.

8. Project future stock conditions and develop rebuilding schedules if warranted, including estimated generation time. Develop stock yield projections in both biomass and numbers of fish in accordance with the following:
   A) If stock is overfished:
      F=0, FCurrent, FMSY, FOY
      F=FRebuild (max that permits rebuild in allowed time)
   B) If stock is undergoing overfishing:
      F= FCurrent, FMSY, FOY
   C) If stock is neither overfished nor undergoing overfishing:
      F= FCurrent, FMSY, FOY
   D) If data limitations preclude classic projections (i.e. A, B, C above), explore alternate models to provide management advice

   Accomplished.

9. Provide a probability distribution function for the base model, or a combination of models that represent alternate states of nature, presented for review.
   - Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments for use with the Tier 1 ABC control rule
   - Provide justification for the weightings used in producing combinations of models

   The Assessment Workshop Report noted that ten sensitivity runs had been considered, one of which had been subjected to stochastic projection. The Assessment Workshop Report advised that "probability distribution functions will be developed for the subset of model recommended by the SEDAR AP for projections … and made available to the Scientific and Statistical Committee (SSC) for the development of management advice, including OFL and ABC". No information relating to these probability distribution functions was presented in the Report.

10. Provide recommendations for future research and data collection. Be as specific as possible in describing sampling design and intensity, and emphasize items which will improve assessment capabilities and reliability. Recommend the interval and type for the next assessment.

Attention was directed to the research recommendations that were made in the Data Workshop Report. The Workshop Assessment Report identified gaps in data, which, if addressed, would improve the assessment capabilities and reliability. Specific sampling design and intensity were not discussed. No recommendations relating to the interval and type for the next Assessment were made by the Assessment Workshop

11. Prepare a spreadsheet containing all model parameter estimates and all relevant population information resulting from model estimates and projection and simulation exercises. Include all data included in assessment report tables and all data that support assessment workshop figures.

    A spreadsheet was not provided in the documentation that was circulated to the Review Panel. The Assessment Workshop addressed this Term of Reference in its Report by providing a table listing the estimates for all parameters used in the model and presenting a listing of each of the input files required to run the Stock Synthesis model for Gulf of Mexico Spanish mackerel.

12. Complete the Assessment Workshop Report (Section III: SEDAR Stock Assessment Report).

    Accomplished.

ToR 18. Make any additional recommendations or prioritizations warranted.

- **Clearly denote research and monitoring needs that could improve the reliability of future assessments**

A number of research needs, which are listed below in priority order, were identified in the course of the desk review. As expected, these were highly consistent with, and thus overlap, a number of the research needs that had been identified by the Data and Assessment workshops.

1. Review or establish programs to collect data on the length composition and age-at-length compositions of landings and discards from each commercial gear and from each recreational fishing mode, and of bycatch of Spanish mackerel from the shrimp fishery. Ensure that the statistical design and spatial coverage of survey or sampling programs are appropriate and that survey or sampling intensity is sufficient to produce estimates of the required precision for the Gulf of Mexico stock of Spanish mackerel. Set goals for performance and establish and monitor performance criteria to assess the quality and completeness of data collection programs. This research need is of the highest priority as it will provide information required by Stock Synthesis to determine the selectivity and retention curves for Spanish mackerel for the commercial, recreational, and shrimp fisheries, the lack of which is a key source of uncertainty in the model.
2. Undertake research to determine reliable relationships between the proportion of females that are mature and both length and age for the Gulf of Mexico stock of Spanish mackerel. This is also of high priority, as the maturity information that is currently used is imprecise. The calculation of spawning stock biomass, a crucial

parameter in the calculation of benchmarks and assessment of stock status, should be based on reliable data.

3. Review programs that are used to collect discard data for Spanish mackerel (and data on the bycatch of Spanish mackerel by the shrimp fishery), and refine these programs to ensure that accurate and complete data estimates of the discards (and bycatch) are collected. Ensure that the statistical design and spatial coverage of survey or sampling programs are appropriate and that survey or sampling intensity is sufficient to produce estimates of the required precision. Set goals for performance and establish and monitor performance criteria to assess the quality and completeness of data collection programs. While this research will not produce immediate improvement in the quality of the assessment, it is important that action is taken as soon as possible to improve the accuracy and precision of the data relating to the quantities of fish that are discarded from each of the fisheries, such that, in the future, the time series of discards become more reliable.
4. A comprehensive study of the stock structure of Spanish mackerel should be undertaken, with the following objectives:
    a. to determine stock structure and the areas occupied by each stock;

    and, assuming that the current view that there are two stocks, *i.e.,* a Gulf of Mexico and a South Atlantic stock, is substantiated,

    b. to determine more reliably the boundary between the Gulf of Mexico and South Atlantic stocks or the extent of overlap;
    c. to extend sampling into Mexican waters and thereby determine the southern boundary of the Gulf of Mexico stock;
    d. to ascertain whether, regardless of the time of year, catches of fish may be assigned reliably to either the Gulf of Mexico or South Atlantic stock on the basis of the area in which they are caught.

    As this study will take some time before completion, it has been assigned a lower priority than the previous items. Determination of the southern stock boundary, however, is important to ensure that other removals from the stock are not occurring in Mexican waters, as such removals are not taken into account in the current assessment.
5. Undertake research to determine the discard mortality of Gulf of Mexico Spanish mackerel that are discarded from the catches of each commercial fishing gear or each recreational fishing mode, recognising that such mortality is likely to differ among different categories into which the discarded fish are classified, *e.g*., "alive", "mostly alive", and "mostly dead".
6. In future stock assessments for the Gulf of Mexico stock of Spanish mackerel, explore whether the use of an age-dependent rather than constant *M* results in a significant improvement in fit, considering the Lorenzen and alternative functional forms of the relationship with age and the alternative of estimating the value of the age-dependent *M* at each age (or range of ages).
7. In future stock assessments, explore the sensitivity of the model to the uncertainty of the landings data.
8. As a low research priority, assess whether, in future refinement of the Stock Synthesis model, sexually dimorphic growth should be introduced. Note that the benefit of this might only be realised if appropriate sex composition data for landings and discards are

available for input, and length and age-at-length compositions are sexually disaggregated.

5. Conclusions and recommendations

After considering the information relating to stock structure, the data that were available for the Gulf of Mexico stocks of cobia and Spanish mackerel, and the details of the assessment for each species, the base model that had been proposed by the Assessment Workshop for each assessment was accepted for use in assessing stock status and in projecting the potential yield and likely stock status over the next six years. The results of the accepted base models, which had been developed using the Stock Synthesis 3 framework, suggested that both stocks were currently (in 2011) not overfished and that overfishing was not currently occurring. While the results of the assessment were imprecise, reflecting the quality and nature of the input data, the results of sensitivity runs for each model suggested that the conclusions drawn regarding stock status were likely to be robust to the uncertainty of the base model results.

Although some of the components of the data for the Gulf of Mexico stocks of cobia and Spanish mackerel were limited and/or uncertain, the datasets that had been collated by the Data Workshops represented the best data currently available for those stocks and appeared adequate for use in assessing, albeit imprecisely, the condition of the two stocks. The models that were developed within Stock Synthesis using these datasets were of appropriate structure and were of a standard that would be considered "best practice" given the types and quality of the data that were available. The explorations of uncertainty and decisions made in the assessments were appropriate. The advice regarding the condition of each stock, *i.e.*, that it is not overfished and overfishing is not occurring, appears sound.

Improvement of the assessments will require the collection of adequate and appropriate data sufficient to characterize the length and age-at-length compositions of catches and discards from both the commercial and recreational fisheries and of bycatches of cobia and Spanish mackerel by the shrimp fishery. These data are essential if selectivity and retention curves are to be accurately determined within the assessment models. Reliable data on maturity are also essential if reliable estimates of spawning stock biomass are to be calculated by the models. Further improvement of the models will require the collection of discard and bycatch data of higher quality from the commercial and recreational fisheries and from the shrimp fishery, and determination of the southern boundaries of both the Gulf of Mexico stocks of cobia and Spanish mackerel.

6. References

Beddington, J. R., and Cooke, J. G. 1983. The potential yield of fish stocks. *FAO Fish. Tech. Pap.*, 242: 1–47.

Diaz, G. A., Porch, C. E., and Ortiz, M. 2004. Growth models for red snapper in US Gulf of Mexico Waters estimated from landings with minimum size limit restrictions. NMFS/SFD Contribution SFD-2004-038. SEDAR7-AW1.

Francis, R. I. C. C. (1993). Monte Carlo evaluation of risks for biological reference points used in New Zealand fishery assessments. *Canadian Special Publication of Fisheries and Aquatic Sciences* **120**: 221-230.

Francis, R. I. C. C. 2011. Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences,* **68**: 1124-1138.

Francis, R. I. C. C. 2012. Report on the 2012 assessments of yellowtail flounder and herring at Woods Hole. National Institute of Water & Atmospheric Research Ltd., Wellington, New Zealand.

Froese, R. and Binohlan, C. 2000. Empirical relationships to estimate asymptotic length, length at first maturity and length at maximum yield per recruit in fishes, with a simple method to evaluate length frequency data. *Journal of Fish Biology* **56**:758-773.

Froese, R., and Pauly, D. (Editors). 2012. FishBase. World Wide Web electronic publication. www.fishbase.org, version (12/2012).

Hoenig, J. M. 1983. Empirical use of longevity data to estimate mortality rates. *Fishery Bulletin*, **82**: 898-903.

Lee, H-H., Maunder, M. N., Piner, K. R., and Methot, R. D. (2012). Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models? *Fisheries Research*, **125–126**: 254– 261.

Lee, Hua-hui, Cheng-Yu Yang, R., and Taylor, I. 2012. User manual for Fishery Simulation Tool. Joint Institute for Marine and Atmospheric Research, University of Hawaii, Pacific Islands Fisheries Science Center, NOAA Fisheries Honolulu, HI USA, 7pp.

Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. *Journal of Fish Biology*, **49**: 627-642.

MacCall, A. 2011. What is the precision of *M*? Abstract *In*: Brodziak, J., J. Ianelli, K. Lorenzen, and R.D. Methot Jr. (eds). 2011. Estimating natural mortality in stock assessment applications. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-F/SPO-119, 38 p.

Maunder, M. N. and Deriso, R. B. 2003. Estimation of recruitment in catch-at-age models. *Can. J. Fish. Aquat. Sci*., **60**: 1204-1216.

Mertz, G., and Myers, R. A. 1996. Influence of fecundity on recruitment variability of marine fish. *Can. J. Fish. Aquat. Sci.*, **53**: 1618 - 1625.

Methot, R.D. 2011. User manual for Stock Synthesis: model version 3.23b. Nov 7, 2011. NOAA Fisheries Service, Seattle, WA.

Methot Jr., R. D., and Wetzel, C. R., 2012. Stock Synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fish. Res.* (in press).

Smith, A. D. M., and Punt, A. E. 1998. Stock assessment of gemfish (*Rexea solandri*) in eastern Australia using maximum likelihood and Bayesian methods. *In* Fisheries stock assessment models. Edited by T. J. Quinn II, F. Funk, J. Heifetz, J. N. Ianelli, J. E. Powers, J. F. Schweigert, P. J. Sullivan, and C-I. Zhang. Alaska Sea Grant College Program, AK-SG-98-01. pp. 245–286.

Appendix 1: Bibliography of all material provided

SEDAR 28 - Gulf and South Atlantic -- Spanish Mackerel and Cobia
Workshop Document List

Document #	Title	Authors
	Data and Assessment Workshop Reports considered in CIE Desktop Review	
	SEDAR 28 – Gulf of Mexico Cobia – Data Workshop Report – May 2012	
	SEDAR 28 – Gulf of Mexico Spanish Mackerel – Data Workshop Report – May 2012	
	SEDAR 28 – Gulf of Mexico Cobia – Assessment Process Report – December 2012	
	SEDAR 28 – Gulf of Mexico Spanish – Mackerel Assessment Workshop Report – December 2012	
	Documents Prepared for the Data Workshop	
SEDAR28-DW01	Cobia preliminary data analyses – US Atlantic and GOM genetic population structure	Darden 2012
SEDAR28-DW02	South Carolina experimental stocking of cobia *Rachycentron canadum*	Denson 2012
SEDAR28-DW03	Spanish Mackerel and Cobia Abundance Indices from SEAMAP Groundfish Surveys in the Northern Gulf of Mexico	Pollack and Ingram, 2012
SEDAR28-DW04	Calculated discards of Spanish mackerel and cobia from commercial fishing vessels in the Gulf of Mexico and US South Atlantic	K. McCarthy
SEDAR28-DW05	Evaluation of cobia movement and distribution using tagging data from the Gulf of Mexico and South Atlantic coast of the United States	M. Perkinson and M. Denson 2012
SEDAR28-DW06	Methods for Estimating Shrimp Bycatch of Gulf of Mexico Spanish Mackerel and Cobia	B. Linton 2012
SEDAR28-DW07	Size Frequency Distribution of Spanish Mackerel from Dockside Sampling of Recreational and Commercial Landings in the Gulf of Mexico 1981-2011	N.Cummings, J. Isely
SEDAR28-DW08	Size Frequency Distribution of Cobia from Dockside Sampling of Recreational and Commercial Landings in the Gulf of Mexico 1986-2011	J. Isely and N. Cummings
SEDAR28-DW09	Texas Parks and Wildlife Catch Per unit of Effort Abundance Information for Spanish mackerel	N. Cummings, J. Isely

Document #	Title	Authors
SEDAR28-DW10	Texas Parks and Wildlife Catch Per unit of Effort Abundance Information for cobia	J. Isely, N. Cummings
SEDAR28-DW11	Size Frequency Distribution of Cobia and Spanish Mackerel from the Galveston, Texas, Reef Fish Observer Program 2006-2011	J Isely and N Cummings
SEDAR28-DW12	Estimated conversion factors for calibrating MRFSS charterboat landings and effort estimates for the South Atlantic and Gulf of Mexico in 1981-1985 with For Hire Survey estimates with application to Spanish mackerel and cobia landings	V. Matter, N Cummings, J Isely, K Brennen, and K Fitzpatrick
SEDAR28-DW13	Constituent based tagging of cobia in the Atlantic and Gulf of Mexico waters	E. Orbesen
SEDAR28-DW14	Recreational Survey Data for Spanish Mackerel and Cobia in the Atlantic and the Gulf of Mexico from the MRFSS and TPWD Surveys	V. Matter
SEDAR28-DW15	Commercial Vertical Line and Gillnet Vessel Standardized Catch Rates of Spanish Mackerel in the US Gulf of Mexico, 1998-2010	N. Baertlein, K. McCarthy
SEDAR28-DW16	Commercial Vertical Line Vessel Standardized Catch Rates of Cobia in the US Gulf of Mexico, 1993-2010	K. McCarthy
SEDAR28-DW17	Standardized Catch Rates of Spanish Mackerel from Commercial Handline, Trolling and Gillnet Fishing Vessels in the US South Atlantic, 1998-2010	K. McCarthy
SEDAR28-DW18	Standardized catch rates of cobia from commercial handline and trolling fishing vessels in the US South Atlantic, 1993-2010	K. McCarthy
SEDAR28-DW19	MRFSS Index for Atlantic Spanish mackerel and cobia	Drew et al.
SEDAR28-DW20	Preliminary standardized catch rates of Southeast US Atlantic cobia (*Rachycentron canadum*) from headboat data	NMFS Beaufort
SEDAR28-DW21	Spanish mackerel preliminary data summary: SEAMAP-SA Coastal Survey	Boylan and Webster
SEDAR28-DW22	Recreational indices for cobia and Spanish mackerel in the Gulf of Mexico	Bryan and Saul
SEDAR28-DW23	A review of Gulf of Mexico and Atlantic Spanish mackerel (*Scomberomorus maculatus*) age data, 1987-2011, from the Panama City Laboratory, Southeast Fisheries Science Center, NOAA Fisheries Service	Palmer, DeVries, and Fioramonti

Document #	Title	Authors
SEDAR28-DW24	SCDNR Charterboat Logbook Program Data, 1993 - 2010	Errigo, Hiltz, and Byrd
SEDAR28-DW25	South Carolina Department of Natural Resources State Finfish Survey (SFS)	Hiltz and Byrd
SEDAR28-DW26	Cobia bycatch on the VIMS elasmobranch longline survey:1989-2011	Parsons et al.
	Documents Prepared for the Assessment Workshop	
SEDAR28-AW01	Florida Trip Tickets	S. Brown
SEDAR28-AW02	SEDAR 28 Spanish mackerel bycatch estimates from US Atlantic coast shrimp trawls	NMFS Beaufort
	Documents Prepared for the Review Workshop	
SEDAR28-RW01	The Beaufort Assessment Model (BAM) with application to cobia: mathematical description, implementation details, and computer code	Craig
SEDAR28-RW02	Development and diagnostics of the Beaufort assessment model applied to Cobia	Craig
SEDAR28-RW03	The Beaufort Assessment Model (BAM) with application to Spanish mackerel: mathematical description, implementation details, and computer code	Andrews
SEDAR28-RW04	Development and diagnostics of the Beaufort assessment model applied to Spanish mackerel	Andrews
	Final Assessment Reports (Not available at time of desktop review)	
SEDAR28-SAR1	Assessment of Spanish mackerel in the US South Atlantic	To be prepared by SEDAR 28
SEDAR28-SAR2	Assessment of Spanish mackerel in the US Gulf of Mexico	To be prepared by SEDAR 28
SEDAR28-SAR3	Assessment of cobia in the US South Atlantic	To be prepared by SEDAR 28
SEDAR28-SAR4	Assessment of cobia in the US Gulf of Mexico	To be prepared by SEDAR 28
	Reference Documents	
SEDAR28-RD01	List of documents and working papers for SEDAR17 (South Atlantic Spanish mackerel) – all documents available on the SEDAR website	SEDAR 17
SEDAR28-RD02	2003 Report of the mackerel Stock Assessment Panel	GMFMC and SAFMC, 2003
SEDAR28-RD03	Assessment of cobia, *Rachycentron canadum*, in the waters of the U.S. Gulf of Mexico	Williams, 2001
SEDAR28-RD04	Biological-statistical census of the species entering fisheries in the Cape Canaveral area	Anderson and Gehringer, 1965

Document #	Title	Authors
SEDAR28-RD05	A survey of offshore fishing in Florida	Moe 1963
SEDAR28-RD06	Age, growth, maturity, and spawning of Spanish mackerel, *Scomberomorus maculates* (Mitchill), from the Atlantic Coast of the southeastern United States	Schmidt et al. 1993
SEDAR28-RD07	Omnibus amendment to the Interstate Fishery Management Plans for Spanish mackerel, spot, and spotted seatrout	ASMFC 2011
SEDAR28-RD08	Life history of Cobia, *Rachycentron canadum* (Osteichthyes: Rachycentridae), in North Carolina waters	Smith 1995
SEDAR28-RD09	Population genetics of cobia *Rachycentron canadum*: Management implications along the Southeastern US coast	Darden et al, 2012
SEDAR28-RD10	Inshore spawning of cobia (*Rachycentron canadum*) in South Carolina	Lefebvre and Denson, 2012
SEDAR28-RD11	A review of age, growth, and reproduction of cobia *Rachycentron canadum*, from US water of the Gulf of Mexico and Atlantic ocean	Franks and Brown-Peterson, 2002
SEDAR28-RD12	An assessment of cobia in Southeast US waters	Thompson 1995
SEDAR28-RD13	Reproductive biology of cobia, *Rachycentron canadum*, from coastal waters of the southern United States	Brown-Peterson et al. 2001
SEDAR28-RD14	Larval development, distribution, and ecology of cobia *Rachycentron canadum* (Family: Rachycentridae) in the northern Gulf of Mexico	Ditty and Shaw 1992
SEDAR28-RD15	Age and growth of cobia, *Rachycentron canadum*, from the northeastern Gulf of Mexico	Franks et al 1999
SEDAR28-RD16	Age and growth of Spanish mackerel, *Scomberomorus maculates*, in the Chesapeake Bay region	Gaichas, 1997
SEDAR28-RD17	Status of the South Carolina fisheries for cobia	Hammond, 2001
SEDAR28-RD18	Age, growth and fecundity of the cobia, *Rachycentron canadum*, from Chesapeake Bay and adjacent Mid-Atlantic waters	Richards 1967
SEDAR28-RD19	Cobia (*Rachycentron canadum*) tagging within Cheasapeake Bay and updating of growth equations	Richards 1977
SEDAR28-RD20	Synopsis of biological data on the cobia *Rachycentron canadum* (Pisces: Rachycentridae)	Shaffer and Nakamura 1989
SEDAR28-RD21	South Carolina marine game fish tagging program 1978-2009	Wiggers, 2010

Document #	Title	Authors
SEDAR28-RD22	Cobia (*Rachycentron canadum*), amberjack (*Seriola dumerili*), and dolphin (*Coryphaena hipurus*) migration and life history study off the southwest coast of Florida	MARFIN 1992
SEDAR28-RD23	Sport fish tag and release in Mississippi coastal water and the adjacent Gulf of Mexico	Hendon and Franks 2010
SEDAR28-RD24	VMRC Cobia otolith preparation protocol	VMRC
SEDAR28-RD25	VMRC Cobia otolith ageing protocol	VMRC
SEDAR28-RD26	Age, growth, and reproductive biology of greater amberjack and cobia from Louisiana waters	Thompson et al. 1991
SEDAR28-RD27	Gonadal maturation in the cobia, *Rachycentron canadum*, from the northcentral Gulf of Mexico	Lotz et al. 1996
SEDAR28-RD28	Cobia (*Rachycentron canadum*) stock assessment study in the Gulf of Mexico and in the South Atlantic	Burns et al. 1998
SEDAR28-RD29	Total mortality estimates for Spanish mackerel captured in the Gulf of Mexico commercial and recreational fisheries 1983 to 2011	Bryan 2012

Appendix 2: Copy of the CIE Statement of Work

Attachment A: Statement of Work for Dr. Norm Hall

Amended Statement of Work

External Independent Peer Review by the Center for Independent Experts

SEDAR 28: Gulf of Mexico Cobia and Spanish Mackerel Assessment Desk 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 Representative (COR), 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 28 will be a compilation of data, an assessment of the stocks, and an assessment review conducted for Gulf of Mexico Spanish mackerel and cobia. The CIE peer review is ultimately responsible for ensuring that the best possible assessment has been provided through the SEDAR process. The stocks assessed through SEDAR 28 are within the jurisdiction of the Gulf of Mexico Fisheries Management Councils and states in the Gulf of Mexico region. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements for CIE Reviewers: Three CIE reviewers shall have the necessary qualifications to complete an impartial and independent peer review in accordance with the statement of work (SoW) tasks and terms of reference (ToRs) specified herein. The CIE reviewers shall have expertise in stock assessment, statistics, fisheries science, and marine biology sufficient to complete the tasks of the peer-review described herein. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall participate and conduct an independent peer review as a desk review, therefore travel will not be required.

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 contact information to the COR, 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 assessment and other pertinent background documents for the peer review. Any changes to the SoW or ToRs must be made through the COR 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.

Desk Review: 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 shall not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COR and CIE Lead Coordinator. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review 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.

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) Conduct an impartial and independent peer review in accordance with the tasks and ToRs specified herein, and each ToRs must be addressed (**Annex 2**).
3) No later than January 25, 2013, 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 Dr. 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.

21 December 2012	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
2 January 2013	NMFS Project Contact sends the CIE Reviewers the assessment report and background documents
9-24 January 2013	Each reviewer conducts an independent peer review as a desk review
25 January 2013	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
8 February 2013	CIE submits CIE independent peer review reports to the COR
15 February 2013	The COR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on changes. The COR 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 COR 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 COR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) The CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) The CIE report shall address each ToR as specified in **Annex 2**,
(3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

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

Support Personnel:

William Michaels, Program Manager, COR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-427-8155

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:

Ryan Rindone, SEDAR Coordinator
2203 N. Lois Avenue, Suite 1100
Tampa, FL 33607
Ryan.Rindone@gulfcouncil.org Phone: 813-348-1630

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.

   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

Annex 2a – Terms of Reference for

SEDAR 28: Gulf of Mexico Cobia Assessment Desk Review

1. Evaluate the quality and applicability of data used in the assessment.
2. Evaluate the quality and applicability 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. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.
5. Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.
6. Evaluate the quality and applicability 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
    - If there are significant changes to the base model, or to the choice of alternate states of nature, then provide a probability distribution function for the base model, or a combination of models that represent alternative states of nature, presented for review.
        - Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments
        - Provide justification for the weightings used in producing the combinations of models
7. If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.
8. Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.
9. Make any additional recommendations or prioritizations warranted.
    - Clearly denote research and monitoring needs that could improve the reliability of future assessments

Table 1. Required MSRA Evaluations for cobia assessment:

Criteria	Definition* (2001)	Current Value* (2001)
Mortality Rate Criteria		
F_{MSY}	F_{MSY}	0.34
MFMT	F_{MSY}	0.34
F_{OY}	75% of F_{MSY}	0.26
$F_{CURRENT}$	F_{2000}	0.30
$F_{CURRENT}$/ F_{MSY}	Percentage of $F_{Current}/F_{MSY}$ > MFMT	0.40
Base M		0.30
Biomass Criteria		
SSB_{MSY}	Equilibrium SSB_{MSY} @ F_{MSY}	3.02 mp
MSST	(1-M)*SSB_{MSY}: M=0.30	2.11 mp
$SSB_{CURRENT}$	SSB_{2000}	
$SSB_{CURRENT}$/ SSB_{MSY}	Percentage of $SSB_{Current}/SSB_{MSY}$ < MSST	0.30
Equilibrium MSY	Equilibrium Yield @ F_{MSY}	1.50 mp
Equilibrium OY	Equilibrium Yield @ F_{OY}	1.45 mp
OFL	Annual Yield @ MFMT	
	2013	
	2014	
	2015	
	2016	
	2017	
	2018	
Annual OY**	Annual Yield @ F_{OY}	
	2013	
	2014	
	2015	
	2016	
	2017	
	2018	

*Definitions and values are subject to change as per guidance from this assessment.

**Based upon current definitions of OY, where OY = 75% of F_{MSY}

Table 2. Projection Scenario Details for cobia assessment

2.1 Initial Assumptions:

OPTION	Value
2012 base TAC	TBD
2012 Recruits	TBD by Panel
2012 Selectivity	TBD by Panel
Projection Period	6 yrs (2013-2018)
1^{st} year of change F, Yield	2013

2.2 Scenarios to Evaluate (preliminary, to be modified as appropriate)

1. Landings fixed at 2013 target
2. F_{OY}= 65%, 75%, 85% F_{MSY} (project when OY will be achieved)
3. F_{MSY}
4. $F_{REBUILD}$ (if necessary)
5. F=0 (if necessary)

2.3 Output values

1. Landings
2. Discards (including dead discards)
3. Exploitation
4. F/F_{MSY}
5. B/B_{MSY}

Annex 2b – Terms of Reference for

SEDAR 28: Gulf of Mexico Spanish Mackerel Assessment Desk Review

10. Evaluate the quality and applicability of data used in the assessment.
11. Evaluate the quality and applicability of methods used to assess the stock.
12. Recommend appropriate estimates of stock abundance, biomass, and exploitation.
13. Evaluate the methods used to estimate population benchmarks and management parameters. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.
14. Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.
15. Evaluate the quality and applicability 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
    - If there are significant changes to the base model, or to the choice of alternate states of nature, then provide a probability distribution function for the base model, or a combination of models that represent alternate states of nature, presented for review.
        - Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments
        - Provide justification for the weightings used in producing the combinations of models
16. If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.
17. Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.
18. Make any additional recommendations or prioritizations warranted.
    - Clearly denote research and monitoring needs that could improve the reliability of future assessments

Table 1. Required MSRA Evaluations for Spanish mackerel assessment:

Note: te = trillion eggs

Criteria	Definition* (as of 2002/2003)	Current Value* (2002/03)
Mortality Rate Criteria		
F_{MSY}	$F_{30\%SPR}$	
MFMT	$F_{30\%SPR}$	
F_{OY}	75% of $F_{30\%SPR}$	0.40
$F_{CURRENT}$	$F_{2002/03}$	
$F_{CURRENT}$/MFMT		0.53
Base M		0.30
Biomass Criteria		
SSB_{MSY}	Equilibrium SSB_{MSY} @ $F_{30\%SPR}$	19.10 te
MSST	(1-M)*SSB_{MSY}: M=0.30	13.40 te
$SSB_{CURRENT}$	SSB_{2003}	17.96 te
$SSB_{CURRENT}$/ MSST		1.34
Equilibrium MSY	Equilibrium Yield @ $F_{30\%SPR}$	8.7 mp
Equilibrium OY	Equil. Yield @ 75% of $F_{30\%SPR}$	8.3 mp
OFL	Annual Yield @ MFMT	
	2013	
	2014	
	2015	
	2016	
	2017	
	2018	
Annual OY**	Annual Yield @ F_{OY}	
	2013	
	2014	
	2015	
	2016	
	2017	
	2018	

*Definitions and values are subject to change as per guidance from this assessment.

**Based upon current definitions of OY, where OY = 75% of F_{MSY}

Table 2. Projection Scenario Details for Spanish mackerel assessment

2.1 Initial Assumptions:

OPTION	Value
2012 base TAC	TBD
2012 Recruits	TBD by Panel
2012 Selectivity	TBD by Panel
Projection Period	6 yrs (2013-2018)
1st year of change F, Yield	2013

2.2 Scenarios to Evaluate (preliminary, to be modified as appropriate)

1. Landings fixed at 2013 target
2. F_{OY}= 65%, 75%, 85% F_{MSY} (project when OY will be achieved)
3. F_{MSY}
4. $F_{REBUILD}$ (if necessary)
5. F=0 (if necessary)

2.3 Output values

1. Landings
2. Discards (including dead discards)
3. Exploitation
4. F/F_{MSY}
5. B/B_{MSY}

Independent Peer Review Report on the SEDAR 28 Desk Review of the Gulf of Mexico Spanish Mackerel and Cobia Assessments

Dr. Beatriz A. Roel

Prepared for

Center for Independent Experts (CIE)

The Centre for Fisheries and Aquaculture Science
Lowestoft Laboratory
Pakefield Road
Lowestoft
Suffolk NR33 0HT
UK

Executive Summary

The assessments of Spanish mackerel and cobia in the Gulf of Mexico were reviewed independently for the Center for Independent Experts (CIE) without consultation with other reviewers or those who produced the assessments. The process extended from 9 January to 4 February 2013. The main conclusions are given separately by species.

The Gulf of Mexico Spanish mackerel stock assessment presented to the SEDAR 28 Assessment Workshop provided output and analysis of results from Stock Synthesis (SS), an integrated statistical catch-at-age model. The model was considered appropriate because it can make best use of the data available including a data-poor historical period. However, data limitations (a recruitment index and data that would inform the model on the stock's response to exploitation) have enforced the requirement for strong assumptions to be made on key parameters.

SS was used to estimate the stock status of Spanish mackerel in the Gulf of Mexico in relation to SPR30% reference points for the Base Run and each alternative model examined. The current stock status was estimated in the Base Run as SSB_{2011} / MSST = 2.96, and exploitation status as $F_{2009\text{-}2011}$ / $F_{SPR30\%}$ = 0.5. Sensitivity tests carried out resulted in estimates of key parameters for management that suggest that the stock is above MSST and exploited below MFMT. The results suggest that the Gulf of Mexico Spanish mackerel stock is not overfished under any of the model scenarios examined and that it is not undergoing overfishing under any of the scenarios examined.

The Spanish mackerel assessment would benefit from the development of an enhanced biological sampling programme. For instance, the development of a research recruitment index would inform the model on the process and possibly preclude the introduction of such strong assumptions.

The Gulf of Mexico cobia assessment was based on results from SS. The assessment used data through 2011 and the time period of the assessment is 1926–2011. Model projections were run from 2013 to 2019. The estimated biomass trajectories showed a sharp decline as the fisheries developed, reaching levels below the minimum stock threshold (MSST) in the late 1980s and early 1990s. Since then the stock appears to have fluctuated above and below the target spawning stock biomass.

Benchmark and reference points for fishing mortality and stock biomass were estimated relative to SPR 30% which were presented for the base case and for each of the sensitivity runs. For cobia, SPR30% reference points are considered valid proxies for MSY. For the base model $F_{current\ (2009–2011)}$ / $F_{SPR30\%}$ was 0.63, whereas the current spawning biomass (2011) relative to MSST was 1.73; on that basis the stock is not considered to be overfished nor undergoing overfishing.

The stock was considered neither overfished nor undergoing overfishing in most of the sensitivity scenarios explored. In the case of low natural mortality, the more pessimistic scenario, both the F_{SPR30} and F_{OY} scenarios led to future stock conditions where the stock was no longer overfished nor undergoing overfishing by 2014. However, fishing under current F predicted a stock undergoing overfishing throughout the projection period.

The lack of information on recruits of age 0 in the data increased the uncertainty of the assessment and the evaluation of the stock relative to reference points. The development of a fishery-independent recruitment index is recommended.

Background

SEDAR 28 consisted of a compilation of data, an assessment of the stocks, and an assessment review conducted for Gulf of Mexico Spanish mackerel and cobia. The Center for Independent Experts (CIE) review for SEDAR 28 was scheduled from 9–24 January 2013, with the deadline for submission of the Peer Review Report on 4 February 2013. The CIE peer review is ultimately responsible for ensuring that the best possible assessment has been provided through the SEDAR process. The stocks assessed through SEDAR 28 are within the jurisdiction of the Gulf of Mexico Fisheries Management Council and states in the Gulf of Mexico region.

Three CIE reviewers with the requisite qualifications to complete an impartial and independent peer review in accordance with the statement of work (SoW) tasks and terms of reference (ToRs) specified herein participated in the process. They were selected on the basis of their expertise in stock assessment, statistics, fisheries science and marine biology being deemed sufficient to complete the tasks of the peer review described herein. Each CIE reviewer participated and conducted an independent peer review as a desk review, so travel was not required.

Description of the Individual Reviewer’s Role in the Review Activities

I participated in all aspects of the review. In particular, I conducted the necessary pre-review preparations, including reviewing background material and reports provided by the NMFS Project Contact in advance of the peer review. I then conducted an impartial and independent (of anyone else) peer review in accordance with the tasks and ToRs specified herein, focusing on the data analyses, parameter estimation and associated uncertainties and the implications for management advice.

SPANISH MACKEREL

Findings by ToR

1. *Evaluate the quality and applicability of data used in the assessment*

A wide range of commercial, recreational and research data was made available for the stock assessment. The data were explored extensively at the Data Workshop (DW).

Life history: The available life history information was reviewed and the main issues were considered carefully. The information does seem to be adequate to conduct a stock assessment. The DW followed the Life History Group recommendation to model the natural mortality rate (M) as a declining Lorenzen function of size consistent with previous SEDAR recommendations.

Discard mortality depends on the conditions of the catching process, including the type of gear utilised. Gillnets had few discards because of its selectivity patterns, but discard mortality does appear to be very high. The shrimp trawl fishery results in very high discard mortality (virtually 100%). There is in fact limited information available on discard mortality for Spanish mackerel, so the values for the gillnet, shrimp trawl and handline fisheries were agreed on the basis of fisher experience and "common sense" and recommended to the Assessment Workshop (AW). Testing the sensitivity to these assumptions would be appropriate here.

The growth models considered seemed to be appropriate, and the decision to combine sexes given practical considerations (the fishery does not distinguish them) is sensible. The scarcity of small fish in the samples did result in growth parameters being rather unrealistic, but the output was adjusted to more biologically reasonable values.

Based on different data sources, it appears that insufficient gonad samples are being collected for histological analyses.

<u>Commercial fishery statistics</u>: Commercial landings data have been developed by gear for the period 1890–2010 and appear to be adequate to support the assessment, although the landings prior to 1950 are considered to be highly uncertain. Landings were aggregated by gillnet, handline and miscellaneous gears, but for assessment purposes, the category miscellaneous is assigned proportionally into gillnet and handline categories.

<u>Shrimp fishery discards</u>: A median value was assumed over the entire period 1945–2011. Initially, this seemed a somewhat questionable decision given that annual shrimp fishery effort was available and a catchability parameter estimated, allowing annual estimates of Spanish mackerel bycatch to be computed. However, bycatch in the shrimp fishery appeared difficult to determine given the low encounter rate between shrimp trawls and Spanish mackerel, and because of irregular observer coverage. As a consequence, the annual variability in shrimp bycatch appeared to be poorly estimated. The decision to impose a super-period based on an estimated mean bycatch seemed therefore to be appropriate.

<u>Commercial discards</u>: These were computed for the period 1998–2010 based on a gear-specific discard rate and effort data. The method seemed to be appropriate but cannot be applied prior to 1998. Discard estimates are, of course, more uncertain than the landings. A weakness here is that the calculated discards may only represent the minimum number of discards made by the commercial fisheries.

<u>Biological sampling</u>: Sample sizes for developing length compositions were inadequate for a considerable number of years and gear strata. This may jeopardise the use of length compositions to correct for potential biases in age compositions in those years.

<u>Recreational fishery statistics</u>: Landings appear to be adequately recorded or estimated for the period covered. For historical recreational landings, a period is defined as pre-1981, with removals for the years 1955–1981 based on a hindcast. It is difficult to assess

the strengths and weaknesses of this data series based on the report of the Data Workshop.

<u>Discards</u>: Sample sizes for Spanish mackerel in the observer data are very small. Some extrapolations were applied and proxies used to calculate the discarded quanta from the different fisheries. There seem to be uncertainties here that need to be reconciled.

<u>Biological sampling</u>: The number of fish sampled is listed, but it is not possible to characterise the sample sizes because the sampling strategy and the targets are not shown. Size data appear to represent the landed catch for the charter and headboat sector adequately. Based on examination of the length composition histograms shown in Fig 4.12.21, sample sizes may have been rather small in recent years.

2. *Evaluate the quality and applicability of methods used to assess the stock.*

The assessment is carried out using Stock Synthesis (SS), a methodology widely used for stock assessment in the United States and elsewhere, including in Europe, where it is used to assess quite a few ICES stocks (ICES, 2012). Interaction with the model developer has contributed to correct implementation of the methodology, and it focused on the handling of discards, which were estimated according to "super periods"; however, the reasons and advantages of using this approach need to be stated more clearly. There is reference to a small CV associated with discards, but it is not clear how that was estimated.

Discard release mortality was incorporated in the model, but the rate seems to be based on rather few data.

A tool to conduct parametric bootstrap analyses was used to characterise uncertainty. This seems to have been a correct decision, because SS provides asymptotic standard errors only, which constitute a minimum estimate.

The model configuration seems to have been appropriate; it includes removals from three directed fisheries:

1. Commercial gillnet (COM-GN)
2. Commercial vertical line gears (Com_RR)
3. Recreational charter, private, headboat and shore anglers (REC)

Of these, the miscellaneous commercial category was apportioned into 1 and 2.

The model fits three indices of abundance (there is some confusion regarding the labelling of the fishery cpue indices on section 3.1.2 of the Assessment Workshop report):

1. Recreational (MRFSS),
2. Commercial line fishery (FWC Vertical line fishery),
3. SEAMAP fishery independent trawl survey.

The indices seem, however, to be very noisy generally, and varying without a trend.

Natural mortality is based on a declining Lorenzen function, and sensitivity to the various assumptions is explored throughout the stock assessment. This is an appropriate procedure because assumptions on the level of M are anticipated to be very influential.

Several parameters were fixed, namely steepness (h) and recruitment variability, but it is not that obvious that the sensitivity to such assumptions was explored sufficiently in the assessment process.

In terms of shrimp fishery discards, a median value was assumed over the entire period 1945–2011. It is not clear why this is done given that annual shrimp fishery effort was available and a catchability parameter estimated, allowing annual estimates of Spanish mackerel bycatch to have been computed.

Model configuration and equations: The shrimp effort index seems to be fitted well by the SS. The index is said to be used to derive annual estimates of F for the shrimp bycatch fleet. This seems to have been done by estimating the catchability Q parameter. However, F is then used to estimate the mackerel bycatch. Figure 3.3c shows the fit (straight line) to the "observed" discards. That procedure is not explained clearly, and specifying the equations would help understanding.

I believe that presentation of the likelihood function would go a long way towards interpreting the model fit to the data.

The fact that the model resulted in an unrealistic estimate for steepness needs further investigation. A plot of the time-series of total landings may provide some insight on the response of the stock to exploitation. Landings between the 1950s and the late 1990s were large, but abundance indices are only available from the 1980s on and do not seem to capture the response of the stock to the decrease in exploitation during recent years. In light of this, fixing steepness to a more realistic value would seem to be appropriate. The value assumed for steepness is the same as that assumed for South Atlantic Spanish mackerel, which would be expected to have similar dynamics.

There are obviously some poor fits to the length composition data, perhaps at least partly related to the model trying to fit the noisy data resulting from small sample sizes. The assessment team chose an assessment model that can make use of all data available, but it is a complex model that requires many assumptions, and the sensitivities to these were not always explored fully. Simpler age-structured production models (Restrepo and Legault 1998; De Oliveira *et al.*, 2007) run from 1981 on would require fewer assumptions, would be less labour-intensive, and may well perform adequately.

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

A number of datasets were examined by the Data Workshop. Those considered appropriate for use in the assessment model were ranked according to their utility as indices of abundance.

1. SEAMAP Groundfish Survey (1987–2011). Recommended for use because it is a long time-series with good geographic coverage.
2. Florida Trip Ticket index (1986–2011) is recommended because it provides good spatial coverage. All indices are based on positive trips only, which is a limitation, and including zero trips would enhance the index's performance as

an indicator of abundance. The handline/trolling index is good because it covers a long period and samples the entire fishery, both inshore and offshore.

3. Recreational MRFSS Index (1981– 2011). This is a Cpue standardised index based on all trips.

The indices proposed are appropriate as indicators of abundance, representing both the commercial and the recreational fisheries as well as providing fishery-independent information. The recreational Headboat Index, based on all trips and standardised by means of a generalized linear model, was not used in the assessment. The reasons behind this decision are not clearly stated in the report.

A shrimp effort index was used to estimate Spanish mackerel mortality in the shrimp fishery.

4. *Evaluate the methods used to estimate population benchmarks and management parameters. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.*

The methods used to estimate population benchmarks and management parameters are based on MSY criteria and yield per recruit. MSY reference points are also supported by ICES, based on international agreements to achieve MSY for exploited stocks by 2015. MSY reference points are based on assumptions about the stock and recruitment functional form that may not be justified by the data. SPR reference points are well accepted proxies for MSY. For precautionary considerations, short-lived species and pelagic stocks should be kept above 30% virgin SPR (Caddy and Agnew, 2004).

The SS estimates of F_REF and SSB_REF (based on 30% SPR) from 1000 bootstrap samples (Figs 3.48-3.49) show that the probability of the stock being outside precautionary levels is very low. Results for the more pessimistic Run 1 also identify the stock as not overfished and not undergoing overfishing. Tables 3.7 and 3.8 provide the necessary values to assess the state of the stock relative to management benchmarks for all configurations presented for review.

5. *Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.*

Deterministic future population status were projected in terms of SSB and SSB and F relative to 30%SPR reference points for two values of steepness (0.8; 0.9) and three levels of exploitation. The projections are not sensitive to the steepness assumed. The results suggest that the stock is projected to remain within safe biological limits given the selected F, and will remain exploited below optimal levels. Note that the top and the bottom panels in Figure 3.52 are the same and that Figure 3.53 was not discussed in the Assessment Workshop report.

Figure 3.53 illustrates future yields for stochastic projections. Yields appear to be stabilising at levels above estimated MSY (Table 3.9).

6. *Evaluate the quality and applicability of methods used to characterize uncertainty in estimated parameters.*
   - *Verify that appropriate measures were provided*
   - *Verify that the implications of uncertainty in technical conclusions are clearly and acceptably stated*
   - *If there are significant changes to the base model, or to the choice of alternate states of nature, then verify that a probability distribution function for the base model, or a combination of models that represent alternate states of nature were provided.*

Asymptotic standard errors were computed for all the parameters estimated. As these tend to underestimate associated uncertainties, the results from a parametric bootstrap procedure (mean and standard error) are presented for key parameters. Mean and standard deviations resulting from bootstrapping were presented. Showing the median as a measure of central tendency and the CVs for comparison between parameters would probably have been a better choice of statistics.

Model estimates are highly sensitive to the value of steepness, which the model estimates poorly. Comparison of the distributions in Figures 3.34 and 3.35 shows that fixing steepness results in more sensible distributions for virgin biomass, SSB ref and R0.

Sensitivity tests were carried out to explore the impact of uncertainties in model parameters such as natural mortality (M) and steepness, data exclusion, data weighting and discard mortality, on parameters that have implications for management. The results from the analyses did not change the perception of the stock relative to reference points because none of the configurations explored suggested that the stock was outside safe biological limits. Interesting to note here is that the alternative exclusion of the abundance indices made little difference to the estimates of key parameters relative to the base run.

7. *If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.*

The stock assessment results are clearly stated in the Stock Assessment report. Table 3.9 addressed the MSRA evaluations requirements. Mortality rate and biomass criteria were estimated for steepness values of 0.8 and 0.9. Annual yields (2013–2022) are provided for F_{MFMT}, F_{OY} and $F_{current}$.

In terms of the requirements for projections, these were all met, although only total yields were provided. Projections were made under three scenarios for fishing mortality: $F_{current}$, F_{SPR30} (F_{msy}) and F_{OY}. Projections under $F_{rebuild}$ or F_0 were not necessary.

8. *Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.*

The SEDAR process results in a rigorous and in-depth review of the data made available and of the assessment. As this is a desk-based review, it lacks any possibility to include interaction with other reviewers of the same material or with the analysts, in my opinion undermining the quality of the review process. Succinctly, questions arising during the review cannot be addressed to those who conducted the analyses, nor was it possible for reviewers of varying skills to complement each others' skills in coming to an overall evaluation of the appropriateness of the methodology or outputs.

9. *Make any additional recommendations or prioritizations warranted.*
   - *Clearly denote research and monitoring needs that could improve the reliability of future assessments*

Increasing sample sizes for the length composition data in both extractions and surveys is recommended if this information is to be used in the assessment. Further, an expanded observer coverage in all Spanish mackerel fisheries would enhance data quality overall.

The sensitivity to uncertainties in the catch data do need to be explored in future.

I agree with the Data Workshop recommendation that there is need of research-based data where Spanish mackerel are caught in sufficiently large numbers to provide a reasonable index of young fish (age 0) abundance. There is currently very little signal of recruitment strength to inform the assessment.

Errata
Assessment Workshop Report
Figure 3.6 caption 2nd line: mackerel commercial **vertical line** gear fishery.
Figure 3.42 upper panel the y-axis needs to be expanded to include all exploitation rate values.
Figure 3.47 define FWC in the figure caption.
Figure 3.49 MFMP definition repeated.

Gulf of Mexico Cobia

Findings by ToR

1. *Evaluate the quality and applicability of data used in the assessment.*

Life history data used in the assessment included natural mortality, growth, maturity and fecundity. There is some uncertainty regarding life history characteristics for this stock because of a general paucity of data, so some common sense decisions were made by the Data Workshop and the Assessment Workshop, such as assuming 50% maturity at age 2 despite recognizing that maturity is better correlated with size. Despite the differential growth of males and females the decision to conduct the stock assessment on the basis of both sexes combined seemed appropriate.

<u>Landings</u>
In terms of commercial landings, the Data Workshop apportioned commercial landings into handline, longline and miscellaneous. For the assessment, commercial landings data (1927– 2011) were aggregated across gears; handline landings represent ~67% of the total commercial landings since 1981. The reason for aggregation is not clearly stated in the workshop reports but presumably is related to inadequate samples sizes for developing length compositions for sufficient year and gear strata, along with inadequate age composition data for all years. Landings data before 1950 are considered to be very uncertain.

Discard estimates have greater uncertainty than the landings and they are likely to be underestimated. The year-specific age structure of cobia could not always be estimated.

The bycatch of cobia in the shrimp fishery was estimated from observer data and SEAMAP trawl data, then scaled using shrimp effort.

Recreational landings data (1950–2011) were aggregated across modes and regions for the assessment. Landings data were collected from 1981 but were hindcast to 1950. Uncertainties in the historical period were estimated, but it is not clear whether those were taken into account in the assessment.

Discard information from recreational fisheries is limited; in other words the discard information reported by anglers cannot be verified, as some surveys simply do not estimate discard levels. Discarded fish size is unknown for all modes covered by MRFSS.

Biological data
Length composition data were collected in both commercial and recreational fisheries with reasonable sample sizes for the recreational fishery. However, given the minimum size limit in operation and the variable growth patterns of cobia, length frequency data did not provide sufficient information on historical recruitment patterns. Age composition data were collected, but there was too little information to be able to track cohorts through time.

Having reviewed the information presented by the Data Workshop and the Assessment Workshop, it was concluded that, despite certain limitations such as those mentioned above, the data provided for assessment were the best available. Every effort had clearly been made to eliminate potential biases and to make the best possible decisions in cases where data were missing. Those decisions and assumptions are fully documented in the report of the Data Workshop.

2. Evaluate the quality and applicability of methods used to assess the stock.

The stock was assessed by means of Stock Synthesis (SS), Methot 2011. Model configurations of increasing complexity were explored, showing that trends in estimated stock biomass remained similar as model complexity increased. The selected model seems to have been appropriate because it allows the assessors to make best use of the information that was available.

The assessment used data through 2011 and the time period of the assessment is 1926–2011. Model projections were run from 2013 to 2019. The assessment was set up to include three fishing fleets and two indices of abundance. The stock was assumed to be at equilibrium at the start of the modelled period in 1926. Removals of cobia were not substantial until after World War II for any of the fisheries.

A single Beverton & Holt stock–recruitment function was estimated in SS, although the reason for selecting this function was not stated. The model was configured to estimate steepness and equilibrium recruitment; however, steepness is very poorly estimated. Variability in recruitment was constrained by fixing sigma R to 0.6. The reality is that there were few data to inform the Beverton & Holt function parameters, and there is concern that the assumptions on steepness may be driving model results. However, the perception of the stock relative to reference points did not change for the range of steepness explored in the sensitivity tests, rendering the assumption at least credible. Estimated parameter standard deviations were generally small and the convergence test results suggested that the model converged with high probability.

Patterns in the residuals from the fit to length frequency data suggest that the model underestimated the numbers of small and large fish in the early period of the commercial data. This is probably related to small sample sizes in which fish at the extremes of the distribution would have been generally under-represented, resulting in selectivity curves that would have driven model predictions for the entire period. Given the paucity of length data, the assumption of time-invariant selection for all fisheries was appropriate. The model seemed to have underestimated small, undersize fish in the recreational fishery, which was hardly surprising.

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

Estimates of SSB, total biomass and fishing mortality were provided by SS. The model predicted the trends in the two indices of catch per unit effort (CPUE) reasonably well, but the uncertainty associated with point estimates appeared to be large. The SSB trajectories show a sharp decline as the fisheries developed, reaching levels below MSST in the late 1980s and early 1990s. Model-predicted SSB is shown with associated 80% asymptotic intervals rather than 90% or 95% confidence intervals, which might be slightly deceiving. Fishing mortality was estimated to have decreased in the early 1990s, and varying with a slightly declining trend thereafter. Whereas F in the recreational fishery has fluctuated quite widely since the late 1990s, fishing mortality in both the commercial fishery and the shrimp fishery declined during the same period. Results from bootstrap analysis show greater uncertainties around the estimated trajectory of F than reflected by 80% asymptotic intervals.

4. Evaluate the methods used to estimate population benchmarks and management parameters. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.

The state of the stock is primarily evaluated relative to 30% spawner-per-recruit population benchmarks. Those seem more appropriate in the case of Gulf of Mexico cobia than MSY reference points, which may be driven by assumptions about the stock–recruit relationship.

Stock status and benchmarks relative to SPR 30% were presented for the base case and each of the sensitivity runs. For the base model $F_{current\ (2009–2011)}$ / $F_{SPR30\%}$ was 0.63, whereas the current spawning biomass (2011) relative to MSST was 1.73; on that basis the stock is not considered to be overfished nor undergoing overfishing. Based on results from the bootstrap analysis for the base case, the $F_{current}$ /$F_{SPR30\%}$ ratio was estimated to be <1, with a high probability, and current SSB /MSST was estimated to be >1, also with a high probability.

The stock was considered neither overfished nor undergoing overfishing in most of the sensitivity scenarios explored. The exceptions were the low M scenario where the stock was considered both overfished and undergoing overfishing, and Run 7; for the latter, only the MRFSS index fitted, which suggested that the stock was overfished.

5. Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.

Model projections carried out with SS were run from 2013 to 2019. The stock was projected under constant fishing mortalities: $F_{current}$, $F_{30\%SPR}$ and F_{OY}. Recruitment was projected by the fitted stock and recruit function. All scenarios explored show an increase in SSB and yields over the projection period as a result of predicting recruitment at a higher level than the recent average. A more pessimistic scenario of future recruitment, e.g., randomly selecting from the estimated recruitment between 2000 and 2009 (omitting 2010 and2011 as highly uncertain), would have been informative.

Fishing at $F_{current}$, $F_{30\%SPR}$ and F_{OY}, the stock is predicted to be within safe biological limits for the base case. For the most pessimistic scenario, low M, the stock is predicted to undergo overfishing under $F_{current}$ but not under $F_{30\%SPR}$ or F_{OY}.

For the base model, under the assumptions made in the projections, fishing the stock at $F_{30\%SPR}$ (F = 0.378) seems to lead to a long-term equilibrium yield below the estimated MSY. Yield per recruit F_{max} is estimated as well above F_{msy}.

6. Evaluate the quality and applicability of methods used to characterize uncertainty in estimated parameters.

- *Verify that appropriate measures were provided*
- *Verify that the implications of uncertainty in technical conclusions are clearly and acceptably stated*
- *If there are significant changes to the base model, or to the choice of alternate states of nature, then verify that a probability distribution function for the base model, or a combination of models that represent alternate states of nature were provided.*

Asymptotic standard errors were computed for all the parameters estimated. As these tend to underestimate associated uncertainties, the results from a parametric bootstrap procedure (mean and standard error) were presented for key parameters. In general, estimates of uncertainty were similar between the two methods. The distributions of F and SSB relative to benchmark parameters from bootstrap samples were shown for the base model, suggesting that there is a high probability that the stock is neither overfished nor undergoing overfishing.

A number of alternative model configurations and states of nature were investigated in sensitivity tests. Iteratively re-weighting the different components did not reveal any conflicting information among alternative data sources. However, this sensitivity run favoured the Headboat index, leading to a conclusion of a slightly more productive stock and experiencing lower fishing mortalities.

The model was only fit assuming a Beverton & Holt stock–recruit relationship but fitting it to an alternative such as a smooth hockey stick would have been informative as a sensitivity test. As a general point, exploring alternative assessment models that do not require strong assumptions on the stock and recruitment functional form would provide clues on the sensitivity of the assessment results to structural assumptions.

Results from the retrospective analysis suggest a stable assessment and show no indication of substantial bias in the assessment. The analysis for age 0 recruits illustrates the uncertainty associated with recruit estimates for the final few years in a given assessment. This is to be expected given the lack of information on recruitment strength for year classes that have not passed through the fishery.

7. If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.

Stock assessment results are accurately presented in the Stock Assessment Report and are consistent with the Panel recommendations.

8. Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.

This review was conducted as a desk review which, in the opinion of this reviewer, might have been undermined by the lack of direct interactions with other members of the Panel and the analysts. The data analyses and stock assessment presented for review were of high standard and state of the art. Terms of Reference were addressed appropriately during the assessment process.

9. Make any additional recommendations or prioritizations warranted.

- *Clearly denote research and monitoring needs that could improve the reliability of future assessments.*

I support the Research Recommendations presented by the Data Workshop. In particular and given the lack of information on cobia recruitment, the development of a recruitment (age 0) index for this important stock is recommended.

A tagging study to identify spawning areas and aggregations would be valuable if additional conservation measures were to be required.

The development of a fishery-independent index of abundance is recommended.

References

Caddy, J. F. and Agnew, D. J. 2004. An overview of recent global experience with recovery plans for depleted marine resources and suggested guidelines for recovery planning. *Reviews in Fish Biology and Fisheries* **14**: 43–112.

De Oliveira, J. A. A., Boyer, H. J, and Kirchner, C. H. 2007. Developing age-structured production models as a basis for management procedure evaluations for Namibian sardine. *Fisheries Research* **85**: 148–158.

ICES. 2012. Report of the Working Group on the Assessment of Southern Shelf Stocks of Hake, Monk and Megrim. ICES Document CM 2012/ACOM: 11.

Restrepo, V. R. and C. M. Legault 1998. A Stochastic Implementation of an Age-Structured Production Model. Alaska Sea Grant College Program • AK-SG-98-01. *Fishery Stock Assessment Models* pp 435 – 450.

Appendix 1: Bibliography of materials provided for review

SEDAR 28 Gulf of Mexico Cobia Data Workshop Report, May 2012

SEDAR 28 Gulf of Mexico Spanish mackerel Data Workshop Report, May 2012

SEDAR 28 Gulf of Mexico Cobia, Assessment Workshop Report, Dec 2012

SEDAR 28 Gulf of Mexico Spanish Mackerel, Assessment Workshop Report, Dec 2012

Working Papers

SEDAR28-AW01 Florida Trip Tickets

SEDAR28-AW02 Spanish mackerel bycatch estimates from US Atlantic coast shrimp trawls

Appendix 2: Statement of Work

Statement of Work

External Independent Peer Review by the Center for Independent Experts

SEDAR 28: Gulf of Mexico Cobia and Spanish Mackerel Assessment Desk 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 Representative (COR), 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 28 will be a compilation of data, an assessment of the stocks, and an assessment review conducted for Gulf of Mexico Spanish mackerel and cobia. The CIE peer review is ultimately responsible for ensuring that the best possible assessment has been provided through the SEDAR process. The stocks assessed through SEDAR 28 are within the jurisdiction of the Gulf of Mexico Fisheries Management Councils and states in the Gulf of Mexico region. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements for CIE Reviewers: Three CIE reviewers shall have the necessary qualifications to complete an impartial and independent peer review in accordance with the statement of work (SoW) tasks and terms of reference (ToRs) specified herein. The CIE reviewers shall have expertise in stock assessment, statistics, fisheries science, and marine biology sufficient to complete the tasks of the peer-review described herein. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall participate and conduct an independent peer review as a desk review, therefore travel will not be required.

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 contact information to the COR, 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 assessment and other pertinent background documents for the peer review. Any changes to the SoW or ToRs must be made through the COR 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.

Desk Review: 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 shall not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COR and CIE Lead Coordinator. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review 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.

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) Conduct an impartial and independent peer review in accordance with the tasks and ToRs specified herein, and each ToRs must be addressed (**Annex 2**).
3) No later than January 25, 2013, 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 Dr. 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.

21 December 2012	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
2 January 2013	NMFS Project Contact sends the CIE Reviewers the assessment report and background documents
9-24 January 2013	Each reviewer conducts an independent peer review as a desk review
25 January 2013	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
8 February 2013	CIE submits CIE independent peer review reports to the COR
15 February 2013	The COR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on changes. The COR 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 COR 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 COR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) The CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) The CIE report shall address each ToR as specified in **Annex 2**,
(3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

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

Support Personnel:

William Michaels, Program Manager, COR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-427-8155

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:

Ryan Rindone, SEDAR Coordinator
2203 N. Lois Avenue, Suite 1100
Tampa, FL 33607
Ryan.Rindone@gulfcouncil.org Phone: 813-348-1630

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.

   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

Annex 2a – Terms of Reference for

SEDAR 28: Gulf of Mexico Cobia Assessment Desk Review

1. Evaluate the quality and applicability of data used in the assessment.
2. Evaluate the quality and applicability 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. Recommend and provide estimated values for appropriate management benchmarks and declarations of stock status for each model run presented for review.
5. Evaluate the quality and applicability of the methods used to project future population status. Recommend appropriate estimates of future stock condition.
6. Evaluate the quality and applicability 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
    - If there are significant changes to the base model, or to the choice of alternate states of nature, then provide a probability distribution function for the base model, or a combination of models that represent alternative states of nature, presented for review.
        - Determine the yield associated with a probability of exceeding OFL at P* values of 30% to 50% in single percentage increments
        - Provide justification for the weightings used in producing the combinations of models
7. If available, ensure that stock assessment results are accurately presented in the Stock Assessment Report and that stated results are consistent with Review Panel recommendations.
8. Evaluate the quality and applicability of the SEDAR Process as applied to the reviewed assessment and identify the degree to which Terms of Reference were addressed during the assessment process.
9. Make any additional recommendations or prioritizations warranted.
    - Clearly denote research and monitoring needs that could improve the reliability of future assessments