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.

16

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)
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}