# 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.

Commercial fishery statistics: 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.

Shrimp fishery discards: 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.

Commercial discards: 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.

Biological sampling: 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.

Recreational fishery statistics: 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.

Discards: 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.

Biological sampling: 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.

Landings
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