SEDAR

Southeast Data, Assessment, and Review

SEDAR 16

South Atlantic and Gulf of Mexico King Mackerel

SECTION V: Review Workshop Report

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

1. INTRODUCTION

1.1. Workshop Time and Place

The SEDAR 16 Review Workshop was held August 4 - 8, 2008 in Jacksonville, Florida.

1.2. Terms of Reference

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

2. Evaluate the adequacy, appropriateness, and application of methods used to assess the stock.
3. Recommend appropriate estimates of stock abundance, biomass, and exploitation.

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

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

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

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

8. Evaluate the SEDAR Process. Identify any Terms of Reference which were inadequately addressed by the Data or Assessment Workshops; identify any additional information or assistance which will improve Review Workshops; suggest improvements or identify aspects requiring clarification.

9. Review the research recommendations provided by the Data and Assessment workshops and make any additional recommendations warranted. Clearly indicate the research and monitoring needs that may appreciably improve the reliability of future assessments. Recommend an appropriate interval for the next assessment.

10. Prepare a Peer Review Consensus Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Complete and submit this report within 3 weeks of workshop conclusion.

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

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

1.3. List of Participants

Reviewers

Guillermo Diaz (Chair) ..... NMFS S&T
Doug Gregory ..... GMFMC SSC/FL Sea Grant
Neil Klaer ..... CIE
Paul Medley ..... CIE
Kenneth Patterson ..... CIE

Presenters

Shannon Cass-Calay ..... NMFS - Miami
Mauricio Ortiz ..... NMFS - Miami
Victor Restrepo ..... NMFS - Miami

Council-Appointed Observers

Brian Cheuvront ..... SAFMC
Ben Hartig ..... SAFMC AP
Albert Jones ..... GMFMC SSC
Anne Lange ..... SAFMC SSC
Robert Muller ..... AW Rep/GMFMC SAP/FL FWC
Donald Waters ..... GMFMC AP
Bob Zales, II ..... GMFMC AP

Council Representation

George Geiger (SAFMC) ..... FL
Michael Ray (GMFMC) ..... TX

Other Observers

Frank Hester..... DSF
Russell Hudson ..... DSF
Tom Ihde..... Univ. of MD
Richard Methot ..... NMFS NWFSC
Mike Wilberg..... CBL

Staff

Patrick Gilles..... NMFS Miami
Rick Leard..... GMFMC
Julie A. Neer ..... SEDAR
Tina O'Hearn ..... GMFMC
Andi Stephens..... SAFMC
Gregg Waugh..... SAFMC

1.4. List of Data Workshop Working Papers

Document #	Title	Authors
Documents Prepared for the Review Workshop		
SEDAR16-RW-01	Virtual Population Analyses of Gulf of Mexico and Atlantic King Mackerel Migratory Groups: Continuity Case and Updated Runs Through July 2008	Cass-Calay, S., M. Ortiz and V.R. Restrepo

Review Panel Consensus

Executive Summary

The assessment was well carried out and used appropriate methods. However, because of uncertainties in stock structure and incomplete data series, a substantial uncertainty in the state of the stock exists. For practical purposes, the most important of these is that it is very uncertain whether good recruitments that appear in some indices means that the available stock biomass of catchable fish in the eastern Gulf will increase in the next years. It will take two to three years for these fish to enter the fishery, at which point an update assessment should be conducted to test whether the expected increase is indeed occurring.

Data

No concerns were raised about US data collection, but the absence of Mexican catch data from the assessment means that the absolute size of the stock can not be estimated. Nevertheless, the assessments contain useful information about trends and relative stock sizes.

It is a problem that few fishery independent surveys cover this stock, and the existing ones are not complete in their spatial or temporal coverage. While much effort has been made to analyze the fishery data to cover for this lack, such analysis cannot be a full and proper substitute for fishery independent survey data concerning a pelagic fish stock. In such stocks, fishery catch rates are often poor estimators of stock abundance

Methods

The methods used are endorsed as the best available and appropriate for the available data. However, a minor correction to the base case was requested by the Review Panel (RP) and this was accepted by the assessment team.

Estimates of Stock Abundance, Biomass, and Exploitation

The uncertainties around the stock assessments due to uncertainties in stock structure and the relationship of the data to the stock are such that considering only base-case assessments would not provide an adequate picture for management purposes. The RP has reviewed a wide range of interpretations of the data and could draw some firm conclusions about the state of the stocks, but other issues remain uncertain. In the face of this uncertainty the RP advocates that estimates be presented in the form of a decision table that illustrates the levels of risk associated with various catch levels.

Population Benchmarks

The RP noted that standard methods had been used to calculate population benchmarks, and did not re-evaluate these methods. Rather, the panel identified what stock status declarations could reliably be made in the light of the uncertainties which had been identified. These declarations are provided.

Both the Gulf of Mexico (GOM) and Atlantic (ATL) spawning stock biomass levels in 2006 were above the MSST, and therefore not overfished. However, it is uncertain whether the GOM stock is currently experiencing overfishing. For the ATL stock, it is uncertain whether overfishing is occurring, but if it is, then this is at a low level.

Methods used to project future population status and characterize uncertainty

The uncertainties in the assessments are so important that they cannot be estimated on the basis of a single assessment with stochastic projections. The RP recommends instead that the results of a number of plausible assessments be projected forwards so that the results can be used for management purposes in the form of a decision table. The Assessment Team has been asked to prepare such tables. The panel also advises on a closer assessment of the assumptions used concerning the shape of the stock-recruitment relationship at low stock sizes.

Presentation of results

Term of reference No 7, "Ensure that stock assessment results are clearly and accurately presented in the Stock Assessment Report, including the Summary Report, and that reported results are consistent with Review Panel recommendations." has not been addressed as the stock assessment report has not been drafted at the time of writing.

Evaluation of the SEDAR process

The Panel strongly recommends that a serious effort be made to fill data gaps (e.g., better designed larval surveys, data to improve stock identification, etc.) and notably to ensure a full coverage of the stock in time and space using methods suited to measuring pelagic fish abundance, such as larval, egg production or acoustic surveys. At present levels of survey effort, the assessment results are unlikely to be precise enough to allow the Management Councils to implement the management procedures currently under discussion (such as setting ABCs for several years in the future on the basis of medium-term projections).

The RP recommended that the behavior of the current control rules be investigated using simulation, to explore whether (and if so, how) the management objectives can be attained using the information available.

The RP had concerns as to the appropriateness of assessing a resource that is apparently migratory and trans-boundary in nature in a national assessment and management structure. This is relevant as the absence of Mexican catch data is a critical source of uncertainty in terms of stock levels and selectivity; better information of the Mexican catch is needed.

Research recommendations

The panel has provided recommendations to help address the concerns noted above and to help improve the accuracy of parameter estimation.

SEDAR 16. South Atlantic and Gulf of Mexico King Mackerel

August, 2008

Review Workshop Terms of Reference

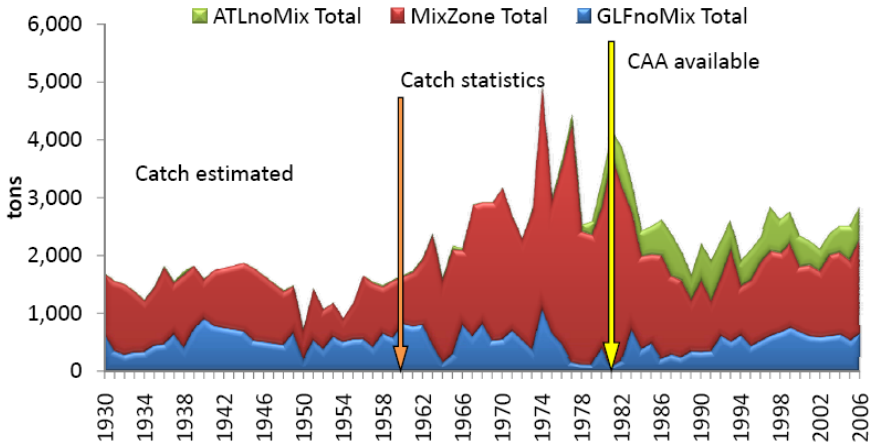
1. **Evaluate the adequacy, appropriateness, and application of data used in the assessment.**

The RP addressed issues of data quality and usage extensively during the meeting, as thoroughly as was possible in the time available and without hands-on experience.

1.1 Landing Data

1.1.1 Commercial

The RP expressed concern that the exclusion of the high reported landings from Mexico may result in an incorrect interpretation of stock status if the GOM and Mexican king mackerel are actually one unit stock. The RP recognized that information on size composition, catch rates, and gears selectivity from that area were lacking and that there were concerns about the quality of the available Mexican landings data (accuracy of landings reports, species identification, etc.).



Except for the available Mexican data, no other concerns were expressed relative to the landings data. The US commercial landings had been accounted for spatially and temporally to include a GOM zone, a mixing zone, and an ATL zone. The mixing zone historically has been the source of about 60% of the total commercial landings.

The RP accepted the AW and DW recommendation that the number of dead discards in the commercial finfish fisheries is considered sufficiently low (about 10-15 thousand fish per year) to be negligible and to not include them in the assessment.

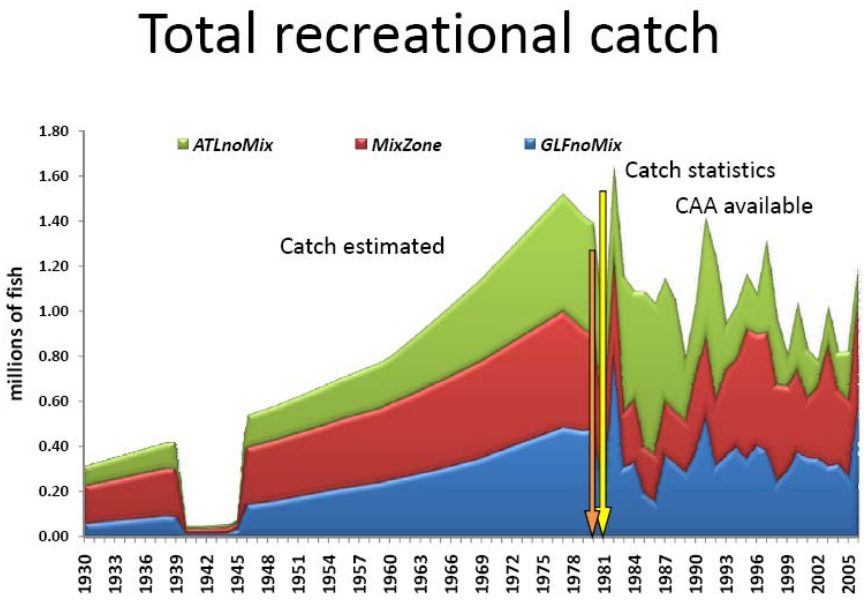
Shrimp Bycatch: The RP agreed with the AW and DW recommendation to exclude shrimp bycatch from the mixing zone in the model since there were few observed occurrences of king mackerel bycatch by shrimp trawlers in this area, and extrapolation

of these using estimated shrimp trawl effort would be highly uncertain. Shrimp bycatch estimates in the GOM were derived from a combination of SEAMAP data and shrimp observer data. The RP also concurred with the AW recommendation to use the delta-lognormal estimation of bycatch in this assessment as an improvement over the standard GLM estimation procedures. However, it should be noted that given the unbalanced nature of the data, the results are sensitive to the estimation procedure used.

1.1.2 Recreational

The RP had no concerns about how the recreational landings were used in the assessment models.

The RP accepted the AW and DW recommendation to apply a 20% release mortality to the MRFSS fishery where fish are released alive and a 33% mortality to the headboat fishery where fish were released both dead and alive with the note that continuity runs do not include discards (B2 portions).



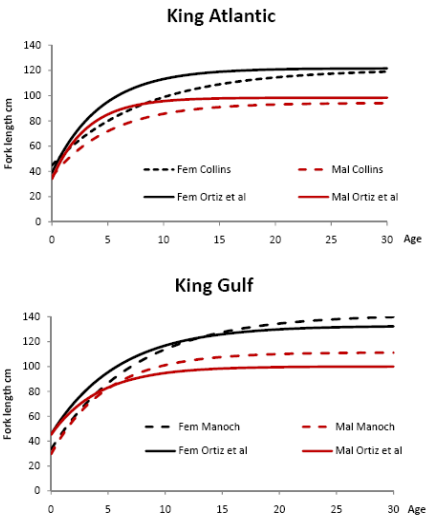
1.1.3 Age Composition

The RP accepted the AW recommendation to use ages 0-11+ for both the GOM and the SA region.

1.2 Life History

1.2.1 Growth

Only age and length data from the non-mixing areas were used for estimating growth curves for the Atlantic and Gulf stocks. Consequently, about a third of the data used in SEDAR 5 was excluded because it was collected in the mixing zone and thus could not be assigned to a particular stock. In addition, new aging data was available and the newly estimated growth functions also took into account minimum size restrictions. As a result, the new models of growth predict faster growth rates for the Atlantic stock and slower growth rates for the Gulf stock than those estimated in SEDAR 5. The RP found the new estimated growth curves to be more appropriate to use in the assessment considering, among other factors, the need to exclude data from the mixing zone.



1.2.2 Stock Composition

The RP accepted the AW recommendation to adopt a 50:50 mixing ratio as the default for the base case. The VPA 2-Box cannot model mixing rates like the SS3 model, so assumptions on mixing ratios had to be made a priori.

It was discussed that there was insufficient data to separate the east and west GOM into two stocks. The DW suggested that a sensitivity analysis could be run excluding all fish west of the Mississippi River. The RP concluded that it would not be instructive to evaluate an eastern Gulf only scenario at this stage. Sensitivity scenarios were run to conduct similar evaluations.

The appropriateness of the level of data aggregation is questionable. While at least two migratory units have been described, over 50% of the fishery is prosecuted on the stocks when they are mixed during the winter. Therefore, separate management of these stock components may not provide the best management advice for king mackerel in the southeastern US unless some mechanism can be developed to identify Gulf and Atlantic biological samples collected in the mixing zone.

It is also possible that a third management unit may also exist in the western GOM. As catches in this area are relatively small, the issue may have relatively little impact, although it should be noted that two of the abundance indices used apply to this region.

1.2.3 Fecundity

The RP accepted the AW and DW recommendations to use the new length-based batch fecundity, a single function for batch fecundity at length for both migratory groups, and the updated fecundity vector based on hydrated oocyte data as reported in SEDAR16-DW-06. However, the RP noticed that the fecundity information was derived from small

samples sizes (32 fish) in the GOM and an effort should be made to estimate new length-based relationship increasing the sample sizes prior to the next assessment.

1.2.4 Maturity

The RP accepted the size/age at maturity information from Finucane *et al.* (1986), but recommends these functions be updated with more recent data.

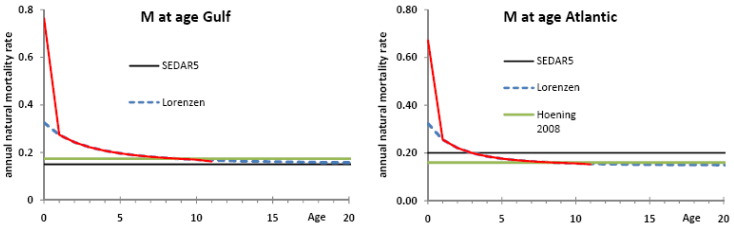
1.2.5 Length-Weight Relationship

The length-weight relationship used in SEDAR 16 differed from SEDAR 5 in that SEDAR 5 used the growth curve to determine the relationship, whereas SEDAR 16 used observed data.

1.2.6 Natural Mortality

The RP accepted the Lorenzen (1996) age-specific estimates of natural mortality (M) scaled to the Hoenig (1983) estimate based on maximum age for king mackerel as presented in the DW report. The RP did not investigate the sensitivity of the assessment results to the assumptions of higher natural mortality at younger ages. The differences seen in the estimates between the Atlantic and Gulf stocks reflect the current observed differences in maximum ages for king mackerel between the ATL (26 years) and the GOM (24 years), which provide Hoenig (1983) estimates of 0.16 and 0.17 year^{-1}, respectively. This procedure resulted in an increase in the Atlantic estimate from the 0.15 used in SEDAR 5 and in a decrease in the 0.20 Gulf estimate used in previous assessments.

- Natural mortality at age
  - Estimates base on Hoening's max age formulation
    - **GOM** 24 yrs M = 0.174 / **ATL** 26 yrs M = 0.160
  - M at age: Lorenzen's formulation [Age 2+ full selected]
  - Adjusted for age 0 .....



1.2.7 Weight at Age

The RP accepted the new weight-at-age estimates recommended by the DW and used by the AW. It was noted that the female weights-at-age used in the VPA2 Box model shows heavier fish at a given age in the GOM than in the ATL resulting in a higher estimated fecundity at age.

1.3 *Indices of Abundance*

1.3.1 North Carolina Trip Ticket Index

The updated index was accepted by the AW panel for use in the VPA2-Box base case for the ATL stock. This index ultimately became the only commercial fishery dependent index in the base model.

1.3.2 Commercial Logbook Index

Because of the complexity of the management regime throughout the last three decades, the AW had difficulty interpreting the fishery dependent indices. The AW had to choose between using either the commercial logbook index or the North Carolina trip ticket index in the SS3 model because that model can accommodate only one index per fishery. Subsequently, this approach of using only one index per fishery sector was carried forward to the VPA analyses.

There was a large difference between the nominal and standardized commercial logbook index in the ATL region. The AW believed that these reflected a change from voluntary to mandatory reporting requirements in 1998. The analysts were not able to entirely remove this influence from the index, so AW group considered using the North Carolina trip ticket index instead of the logbook index. Ultimately, the AW decided to use the logbook data for the GOM no-mixing zone and to use the North Carolina trip ticket index for the SA no-mixing zone. The AW also proposed that the ATL commercial logbook index be used as a sensitivity run for the ATL region.

1.3.3 Marine Recreational Fishery Statistical Survey (MRFSS)

The AW concluded that bag limits did not appear to affect fishing behavior as fishermen frequently exceed the bag limit, and recommended the inclusion of the MRFSS CPUE index in the assessment for both VPA and SS3. There was some concern expressed by the AW over the large variability in the MRFSS index for the ATL. However, the AW determined the MRFSS index to be usable since only recreational fishing trips that were considered to potentially be able to catch king mackerel and only the intercept data were used to develop the index.

1.3.4 Headboat

This index, with the AW recommendation that data collected during closed seasons be excluded, was accepted by the RP as a plausible abundance index.

1.3.5 Fall Plankton Survey (GOM)

Fall plankton survey (also referred to as the SEAMAP ichthyoplankton survey) was used within the VPA model runs as an index of spawning stock biomass (SSB) for the GOM stock. The RP agreed with the decision to include this index particularly given that it was the only fishery independent index used in the assessment for the adult stock.

1.3.6 Shrimp Bycatch Index (GOM)

The RP agreed with the AW that since the shrimp bycatch index is derived from the SEAMAP Groundfish survey data it was not necessary to include it as a second fishery independent recruitment index.

1.3.7 Small Pelagics Trawl Survey (GOM)

The RP agreed with the DW and AW recommendation not to use the small pelagic trawl survey from the GOM, given the very short length of the time series available.

1.3.8 South Atlantic Shark Gill Net Index

The RP agreed with the DW and AW recommendation not to use the South Atlantic shark gill net index because the number of drift gill net vessels in the shark fishery has decreased, few trips were observed each year, the survey only had a small area of coverage, and changes in target species may have occurred . In addition, gill nets only make up a small percentage of the overall king mackerel landings in recent years.

1.3.9 SEAMAP Shallow Water Trawl (ATL)

The DW and AW recommended using the index for mid age-0 king mackerel for both the VPA and SS3 models. The SEAMAP shallow water trawl survey was used as an index of age-1 abundance for the ATL stock under SEDAR 5. However, most of the king mackerel caught during this trawl survey were 40 to 430 mm FL (SEDAR16-DW-9) and the SEDAR 16 DW recommended it as an index for age-0 king mackerel. This is the only index for ages 0 available for the ATL. It was noted that there was a high degree of variability prior to 2001.

1.3.10 SEAMAP Groundfish Survey (GOM)

The AW included the SEAMAP groundfish survey as an index of GOM age-0 abundance. However, the recent four years of increased king mackerel catches in this index were of such a magnitude and had such an influence on model outcomes, the RP was concerned about relying on this sole recruitment index for predicting future population growth, particularly as its extent was limited to the western GOM. Of particular concern was that these much larger year classes had not yet be seen in the catches of the more recent younger ages in the fishery. The RP was also concerned that the three first years of the series used in the VPA had zero values and that these had been replaced by the lowest value in the series. The RP requested sensitivity runs excluding this index, then including index without the first three years of the series.

1.3.11 Summary of data concerns

The lack of Mexican data means that the absolute size of the stock cannot be estimated because an important part of the catches are missing. While this lack should be remedied, useful conclusions about the state of the stock and local management implications can still be made.

The use of fecundity estimates in the estimation of spawning stock biomass is useful only if fecundity is estimated reliably and if it varies substantially with time or with size of fish. These conditions do not pertain at present because the sample size is small, no time-series is available and fecundity appears to be linear with respect to fish weight.

Stock identity is still not reliably described; for example, the affinity of the western Gulf fish to fish in other areas is not known with certainty. The assessment and management system may not be robust to such uncertainties.

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

The RW addressed assessment methodology and interpretation thoroughly.

2.2 Stock Assessment Models

2.2.1 Use of the Stock Synthesis Model

Much of the AW's attention was directed at developing a new area-based model that would describe the population dynamics, migration and exploitation of the relevant stocks of king mackerel. This was perceived, according to the AW's Term of Reference 8, as required for the calculation of management parameters for GOM and ATL migratory units, and follows the recommendation of SEDAR 5.

Such a model was developed using Stock Synthesis 3 (SS3). This model is structured so as better to reflect perceptions of the life history of this stock and to estimate population parameters by maximum-likelihood fitting with respect to the available observations with a minimum of data pre-processing.

The use of the model ran into two difficulties. Firstly, it was not clear that the model could complete its calculations correctly due to hardware and operating system limitations. Secondly, the AW found that the estimates of the population parameters of the two migration groups were very strongly interdependent and could not be estimated separately.

Faced with a perceived need to produce separate management parameters for the two regions, the AW took a decision to base its advice on VPA and to abandon the use of the three-area model.

The RP considered that a conclusion to be drawn from this analysis is that independent assessment of the two migration groups is currently not possible without making arbitrary assumptions and without excluding a substantial amount of biological data. The AW did not adequately follow-up this result, which could have led to a single assessment covering both migration groups. This was because (a) inappropriate terms of reference constrained the analysis (see **Section 2.3**), and (b) data pre-processing had already been completed on a migration-group basis. Due to time constraints, this matter could not be revisited during the RW.

The RP agreed that the SS3 model requires further evaluation and testing before being used in a management context. At this stage the RP concurs with the AW that the SS3 approach was not adequate for the stock assessment, even though this method had substantial theoretical advantages and could lead to better knowledge about the reliability of fish stock assessments.

However, the exploratory use of the SS3 raises serious concerns that the management benchmarks of the two migration units cannot be estimated separately with the available

data. Furthermore, as over 60% of the commercial catches and over 50% of the recreational catches are taken from the stocks when they are mixed, the possibility of assessing and managing the two units separately needs to be questioned further.

The AW decided to base its advice on a conventional two-area "virtual population analysis" (VPA) approach. Given the foregoing concerns, the RP concludes that this decision may not be appropriate.

2.2.2 Continuity Case VPA

The continuity case assessment is intended to show the effects of new observations, while keeping model assumptions as unchanged as possible given the new data. The RP examined the continuity case against the criteria given in **Section 2.4**. The following table summarizes the RP conclusions.

Criterion	Continuity case	Review Panel Conclusion
All relevant data to be included unless there is a clear reason for rejection, no "filled-in" observations to be used.	Yes	Acceptable
Data screening for high residuals and sensitivity	Not tested	Not needed for continuity case
Model screening to test robustness to alternative model structures	Not tested	Not needed for continuity case
Residual pattern screening for trends and appropriate fit	Not screened	Not needed for continuity case
Credibility check on exploitation pattern	Not explicitly assessed	High variability in F at last age leads to doubts on credibility of exploitation pattern in that year.
Credibility check on trends	New assessment shows higher biomass level in ATL in whole time-series and high recruitment in GOM in last three years.	ATL higher biomass level is probably due to new exploitation pattern, which has a much lower F at last age estimated than in SEDAR 5. High recruitments seem due only to high values in the shrimp by-catch index in last few years.
Parameters estimated with reasonable precision	Not described.	Missing
Full documentation of input data	Yes	Good documentation

Criterion	Continuity case	Review Panel Conclusion
Structural model equations	Yes	Does not say if *q*s are conditional or are estimated as free parameters.
Observation error-model equations	Yes	Variance-estimating method is not fully described
Description of estimating algorithm	Reference to standard software	Acceptable
List of final parameter estimates with s.d.	Parameter s.d. and covariances not provided.	Requested by RP
Simulation testing of algorithm	No references made in report, but the method has been simulation tested and the documentation is available at ICCAT.	Acceptable.
Source code and documentation available.	References to program manual provided.	Acceptable

The RP concluded that the continuity case was acceptable and indicated (1) the strong influence of the GOM shrimp bycatch CPUE index in creating a new perception of the state of the stock, and (b) the estimation method may be unstable in estimating selection pattern, and hence historic perceptions of stock size.

2.2.3 Base Case VPAs

The AW's work led to the proposition of base case VPAs which differed from the continuity cases in the following elements:

Element	**SEDAR 5 usage**	**Proposed new usage**	**Review Panel Comment**
Proportion of GOM fish in the mixing zone	Assumed 100%	Assumed 50%	This is reasonably supported by data.
Age-range	Not included	Include age 0 in ATL models	Acceptable
F-parameterizations	F on youngest ages fixed by F ratio from separable VPA.	Estimate more F parameters, with penalty function on change in F at age at sigma =0.4 on ages 3 to 9.	Acceptable
Life history parameters	Use available data on fecundity and growth	Use new data on fecundity and growth	Correct.
Natural Mortality	Fixed M at age, 0.15 for	Use size-related natural	Acceptable.

	ATL and 0.20 for Gulf of Mexico	mortality estimates, but with same average values.	

Following the RP's positive evaluation of the reasons for changes from the continuity case VPAs, the Panel evaluated the assessment against the criteria in **Section 2.4**.

Basic evaluation of base case assessments		
Criterion	**Base case**	**Review Panel Conclusion**
All relevant data to be included unless there is a clear reason for rejection, no "filled-in" observations to be used.	Correction made	Filled-in observations used for early years of SEAMAP survey unacceptable; should be corrected (see section 2.3).
Data screening for high residuals and sensitivity	Not tested	Requested from assessment panel: observed versus expected, residual versus time, QQ plot. This was provided at the meeting.
Model screening to test robustness to alternative model structures	Not tested	See section 2.3
Residual pattern screening for trends and appropriate fit	Screened in informal process but not fully documented.	There are substantial residual trends and index divergences. See section 2.3
Credibility check on exploitation pattern	Not addressed	High variability in F at last age leads to very dome-shaped exploitation pattern. This is considered credible as larger fish are not commercially targeted due to lower value.
Full documentation of input data	Yes	Good documentation
Structural model equations	Yes; Does not say if q's are conditional or are estimated as free parameters.	q's are model parameters
Observation error-model equations	Yes	Variance-estimating method is not fully described.
Description of estimating algorithm	Reference to standard software	Acceptable
List of final parameter estimates and s.d.	Parameter s.d. and covariances not	Requested by RP. Estimates of parameter CV s were acceptable.

Basic evaluation of base case assessments		
	provided in AW report but CV s were made available at the meeting.	
Simulation testing of algorithm	No references made in report, but the method has been simulation tested and the documentation is available at ICCAT.	Acceptable.
Source code and documentation available.	References to program manual provided.	Acceptable

The RP concluded that the preparation and documentation of the assessment base cases was generally of a high standard. However, the use of lowest observations to replace zero observations under assumption of a lognormal distribution was erroneous. The RP requested that the GOM base case be corrected to take this into account. 'Corrected base case', therefore, refers to a VPA run where the original base case was modified by deleting the first 3 years of data of the SEAMAP survey index. The additional elements concerning residual patterns were requested during the RW and are to be provided as addenda to the Assessment Workshop Report.

2.3 Sensitivity testing of the base cases

The RP identified six principal issues that could affect the outcome of the assessment, as below:

Stock Structure The appropriateness of the level of data aggregation is questionable. While at least two migratory units have been described, over 50% of the fishery is prosecuted on the stocks when they are mixed during the winter, so separate management of these stock components may not provide the best management advice. The panel would have liked to test a combined-area assessment, but the structure in which the data had been pre-processed made this impossible at the meeting. An appropriate research recommendation was developed.

The RP considered that the terms of reference Nos. 6 and 8, set to the assessment working group were inappropriate. Fish stock assessments need to be calculated on the basis of functional fishery units taking into account both biological and fleet operation factors. Allocation decisions between management areas should be made outside and after fish stock assessment workshops.

It is also possible that a third management unit exists in the western GOM. Although catches in this area are relatively small, two of the abundance indices used in the

assessment of the entire GOM stock were derived from data in this region. If the stock structure hypothesis is incorrect, large errors in perceptions of stock size are possible.

Exclusion of Mexico catches Tagging data show extensive migrations between the US GOM (and especially the Western Gulf) and Mexican waters, where a fishery also takes place. As these reported catches are of the same order as the US catches, it is necessary to include them in the GOM assessment. As no sampling data were available concerning these catches, only the landings (and not their age or length-structure) could be included. A sensitivity run was provided and shows generally similar trends in biomass and fishing mortality to the base case but at different levels. Fishing mortality is estimated as 10% higher compared to the MFMT while the B_{MSY}, catch forecasts, and MSST are approximately doubled.

Use of fishery-dependent indices Such indices are in many areas considered inappropriate for the assessment of pelagic fish stocks, even though they are used where fishery-independent surveys are not available. The review panel questioned the use of fishery-dependent data series unless they could be shown to reflect stock abundance. The fishery-dependent indices used show surprisingly little correlation among themselves (Figures 2.3.1 and 2.3.2) adding to concerns that they may not all adequately reflect stock abundance. Two sensitivity runs were requested for each area, excluding either the fishery-dependent or the fishery-independent indices. These fits were poor, but showed wide divergence, with fishery-independent indices leading to much higher estimates of stock size (B_{2006}/MSST revised from 1.499 in the corrected base case down to 1.074 based on fishery-dependent indices or up to 2.773 for the fishery-independent indices). Conversely, a much higher fishing mortality is estimated using the fishery-dependent indices (F_{2006}/MFMT=1.477) than for the fishery-independent indices (F_{2006}/MFMT=0.509). For the ATL stock a paucity of data prevented similar comparison between dependent and independent indices (Note that the ATL assessment had only one fishery independent index which corresponded to age 0, therefore running the model with only that index was considered not a plausible alternative).

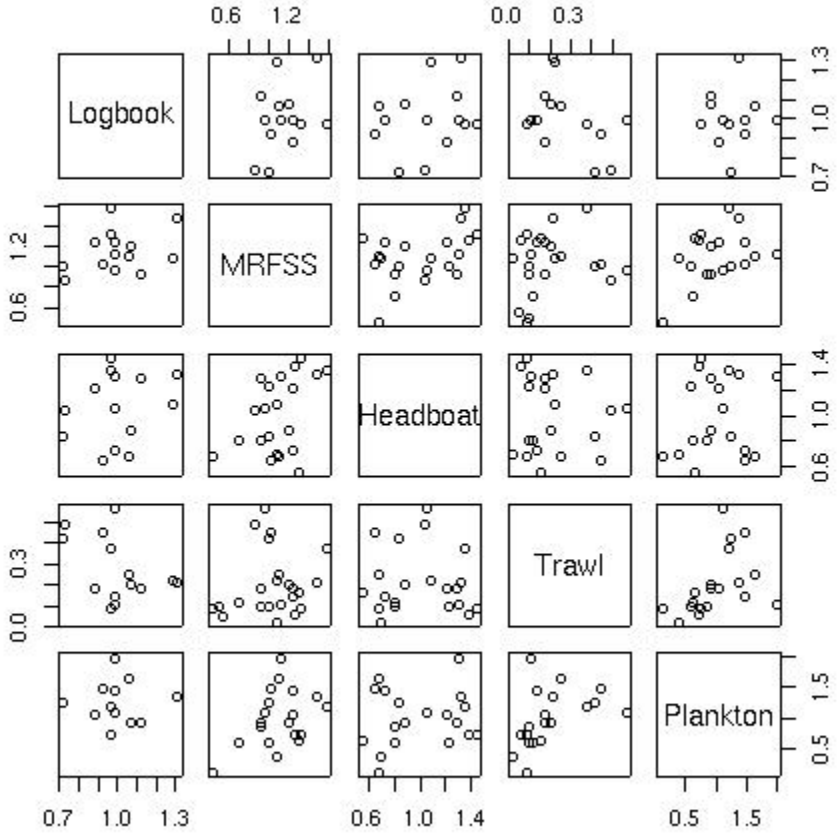


Figure 2.3.1. Pairwise scatterplots of the abundance indices used in fitting the base case for the Gulf of Mexico.

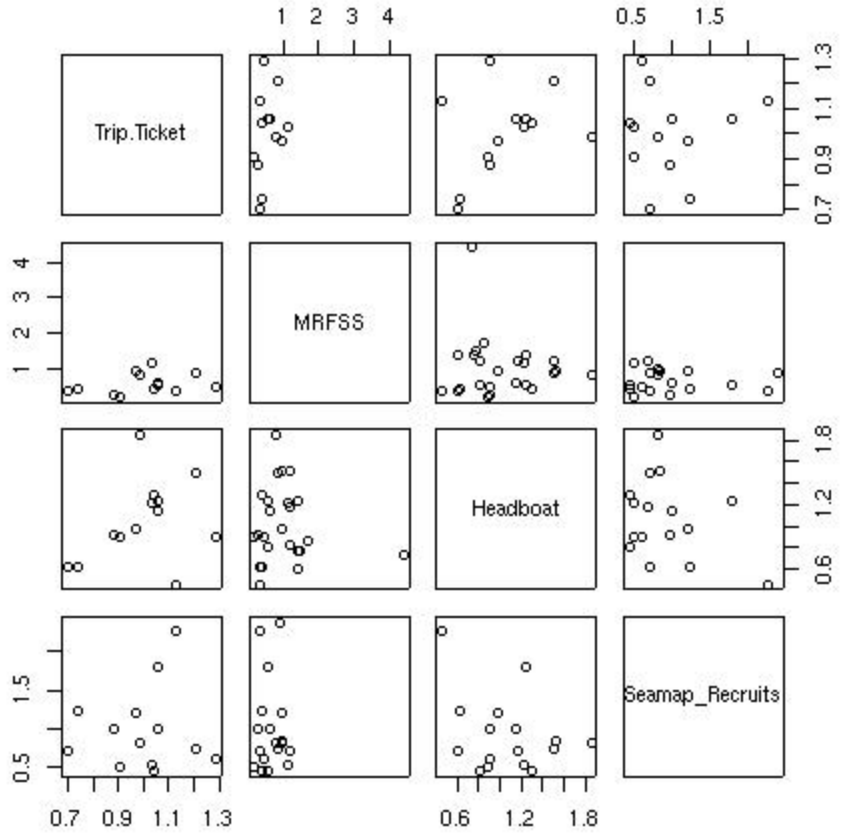


Figure 2.3.2. Pairwise scatterplots of the abundance indices used in fitting the base case for the Atlantic.

Iterative re-weighting Within the range of assessments bracketed by using either fishery-independent or fishery-dependent indices, various solutions can be found according to the statistical weights assigned to the various index series. In the "base case" the indices were assigned equal overall weightings. In principle, these weightings can be estimated within the assessment model. This has the disadvantage of potential numerical instability, but the advantage that the final model fit may better coincide with the index series that appear to be more precisely estimated. The RP requested such an additional run with the purpose of evaluating the sensitivity of the base case to using model-dependent information on index precision. In this case, the assessment model estimated lower variances (and hence closer fits) to the fishery-independent indices.

Exclusion of age 0 survey data in the GOM The observations of very high recruitment in the SEAMAP groundfish surveys in the western GOM have a strong influence on perceptions of stock dynamics. There are four concerns about accepting these estimates at face value:

- the high recruitments seen in the surveys do not appear as highly abundant year-classes in the catch-at-age data;
- the surveys are carried out in the western GOM whereas most of the fishery is deployed in the eastern Gulf, and there is concern that these may not be a single stock unit;
- ecological conditions may have changed in the area following the large reduction in the GOM shrimp fisheries;
- the SEAMAP trawl surveys on which they are based are not executed according to a sufficiently balanced statistical design, which results in a large sensitivity of the index to the method used in estimating inter-annual changes in abundance.

The RP wished to quantify the uncertainty introduced by these concerns by assessing the influence of these data on the assessment, either by excluding the last four years of survey estimates or by excluding the entire 0-group data series. Detailed diagnostics of the sensitivity runs are provided as addenda to the Assessment Workshop report.

Atlantic summer commercial logbook index The review panel also requested to test the sensitivity of the Atlantic assessment to the use of fishery-dependent information in the mixing area in the summer months (April-October) when few Gulf fish should be present. A model fit was calculated by using the commercial fishery for the Atlantic area ("Relative Index" given in Table 11 of SEDAR16-DW-22), and the MRFSS-ATL series (Table 3.5 of the SEDAR 16 Assessment Workshop Report). This showed a 38% reduction in F_{2006}/MFMT and a 23% increase in B_{2006}/MSST. The results of this scenario are presented in Table 2.3.2 under the heading 'New Index'.

The sensitivity of the management parameters to plausible alternative assumptions is summarized in Tables 2.3.1-2.3.2.

Table 2.3.1. Management-related parameters estimated in the base case and in six sensitivity runs for the GOM stock.

	Corrected Base	Include Mexican catches	Only Fisheries Dependent Indices	Only Fisheries Independent Surveys	Survey variances estimated iteratively	Excludes the last 5 years of SEAMAP trawl survey	Excludes Age 0 from analysis
Convergence	Yes	Yes	Yes	Somewhat sensitive to initial estimates of terminal F	Yes	Very sensitive to initial estimates of terminal F	Yes
F30%SPR	0.187	0.210	0.157	0.151	0.193	0.156	0.23
F40%SPR	0.134	0.137	0.116	0.106	0.141	0.108	0.16
0.65*F30%SPR	0.122	0.137	0.102	0.098	0.126	0.101	0.15
0.75*F30%SPR	0.141	0.158	0.118	0.114	0.145	0.117	0.17
0.85*F30%SPR	0.159	0.137	0.102	0.098	0.126	0.101	0.2
Yield equilibrium F30%SPR	10.827	29.189	8.627	9.769	9.802	10.143	7.763
Yield equilibrium F40%SPR	9.972	27.183	7.939	9.273	8.913	9.601	6.855
Yield equilibrium 0.65*F30%SPR	9.434	27.712	7.544	9.118	8.499	9.462	6.557
Yield equilibrium 0.75*F30%SPR	9.947	28.462	7.979	9.420	9.008	9.773	6.994
Yield equilibrium 0.85*F30%SPR	10.335	28.969	8.305	9.617	9.396	9.978	7.350
MSST	2615	6030	2447	2446	2445	2444	1532
Fcurrent	0.155	0.164	0.232	0.077	0.245	0.173	0.200
B2006	3921	11350	2627	6784	2890	2649	3076
Fcurrent/MFMT	0.826	0.779	1.477	0.509	1.268	1.107	0.870
B2006/MSST	1.499	1.883	1.074	2.773	1.182	1.084	2.01

Table 2.3.2. Management-related parameters estimated in the base case and in six sensitivity runs for the ATL stock.

	Base	Only Fisheries Dependent Indices	Only Fisheries Independent Surveys	Survey variances estimated iteratively	New index
Convergence	Yes	Yes	Yes	Yes	Yes
F30%SPR	0.256	0.255	0.262	0.241	0.243
F40%SPR	0.174	0.173	0.178	0.168	0.169
0.65*F30%SPR	0.167	0.166	0.170	0.157	0.158
0.75*F30%SPR	0.192	0.192	0.196	0.181	0.182
0.85*F30%SPR	0.167	0.166	0.170	0.157	0.158
Yield F30%SPR	8.964	8.796	9.001	8.669	9.908
Yield F40%SPR	8.122	8.012	8.131	7.824	8.951
Yield 0.65*F30%SPR	7.996	7.908	8.009	7.610	9.156
Yield 0.75*F30%SPR	8.375	8.265	8.397	8.018	9.557
Yield 0.85*F30%SPR	8.662	8.530	8.691	8.331	9.850
MSST	1827.506	1827.196	1826.675	1826.734	2073.946
Fcurrent	0.258	0.277	0.555	0.148	0.175
B2006	2443.000	2982.000	1064.000	4026.000	3404.000
Fcurrent/MFMT	1.007	1.085	2.121	0.615	0.722
B2006/MSST	1.337	1.632	0.582	2.204	1.641

Conclusions

The RP considered that the assessment was limited because of the absence of Mexican data while tagging information strongly indicates important stock mixing across the area. This is a "straddling" stock, whereby obligations concerning joint research and management exist in the UNCLOS Agreement for the implementation of the provisions of the Convention relating to the conservation and management of straddling fish stocks and highly migratory fish stocks. Although the USA has ratified this agreement, Mexico has not yet done so although joint management and data collection concerning other stocks is already in place.

The sensitivity run including using the available landings information from the Mexican fishery shows the high sensitivity of management parameters related to absolute measures of stock size.

The uncertainty in the assessment due to missing data and to the plausible alternative assessment structures are so large that the RP did not examine the base model parametric uncertainty estimates in detail, nor the medium-term projections, because the uncertainty

in providing management advice is largely due to the variability among alternative model assumptions and specifications.

The RP considered the adequacy and appropriateness of the assessment for various purposes, in the light of structural uncertainty as indicated in the sensitivity runs

Gulf of Mexico Stock		
Purpose	**Adequacy and Appropriateness**	**Reviewer's comments**
Estimation of absolute management benchmarks (MSST, B_{MSY}) related to biomass.	Inadequate. These parameters are very sensitive to the missing Mexican catches which affect the perception of the size of the stocks.	MSST is unknown, but in the range 2444 to 6030 million lbs.
Estimation of stock status with respect to MSST	Adequate.	The stock is above MSST.
Estimation of stock status with respect to MFMT	Inadequate.	The fishing mortality is estimated as between 49% below MFMT and 48% above MFMT.
Evaluation of general trend in stock development	Adequate.	Stock size is increasing and there is an indication of more abundant recent recruitments
Determination of ABC in the short term	Adequate.	Catches corresponding to F30% SPR are fairly robust to model uncertainty in the range 8.63 to 10.27 million lbs, excluding Mexican catches but including shrimp by-catches.

Atlantic Stock		
Purpose	**Adequacy and Appropriateness**	**Reviewer's comments**
Estimation of absolute management benchmarks (MSST, B_{MSY}) related to biomass.	Adequate	MSST in the range 1827 to 2074 million lbs from alternative models. Bootstrap estimates of uncertainty are very tight and do not seem credible.
Estimation of stock status with respect to MSST	Adequate.	The stock is above MSST.
Estimation of stock status with respect to MFMT	Adequate.	The stock is very probably not undergoing overfishing.
Evaluation of general trend in stock development	Adequate.	There are trends showing a decline in biomass and over the time series.
Determination of equilibrium ABC	Adequate.	Catches corresponding to F30% SPR are fairly robust to model uncertainty in the range 0.24 to 0.26 million lbs.
Determination of ABC in the medium term		
Statements of stock status with respect to management benchmarks	Adequate.	

2.4 Evaluation criteria

The RP evaluated the assessment methodology against the following criteria, which where first promulgated at SEDAR 9:

Evaluation criteria for assessments (source: SEDAR 9)

1. All relevant data should be used, unless there is an *a priori* reason to exclude a data series, or a sound *a posteriori* reason can be identified. Data should be real observations, not "filled-in" using assumptions or other criteria, to the extent possible. Fish stock assessment depends on having reasonably long time-series of catch, effort and fishery-independent abundance estimates.

2. Conclusions about stock status with respect to reference points should be robust to underlying assumptions about data and structural model, e.g. reliance on filling-in assumptions, dependence on most contested parts of the data sets.

3. Assessments should include the following:
    - 3.1 Data screening, to check assumptions in 1 and 2.
    - 3.2 Model screening, to see if broadly similar conclusions are drawn from different models, including sensitivity to constraints etc.

3.3 Residual pattern screening: Does the model replicate the trends in the data?
3.4 Credibility check: are the estimated model parameters reasonable (e.g. selection pattern, r, B_0/B_{MSY}, trends in F etc. in the context of biological knowledge about the stock and the fishery?
3.5 Variance estimates (or posteriors) for the estimated interest parameters, and a priori model testing, using simulated data, which should demonstrate that the model has useful precision in predicting interest parameters when presented with data.

4. Assessment documentation should include:
4.1. Data used to fit the assessment model.
4.2. Structural model equations, including process-error model if applicable
4.3. Observation-error model
4.4. Description of estimating algorithm
4.5. List of final parameter estimates and their s.d.s
4.6. Computational validation, including simulation testing
4.7. Source code (and ideally documentation) of the programs used should be made available.

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

The RP accepted the base cases provided by the AW for the GOM and ATL stocks as providing one of several plausible estimates of stock abundance, biomass and exploitation. However, the base cases alone do not provide sufficient information about the uncertainty of these estimates.

Gulf of Mexico stock:

- Stock abundance and biomass: The stock is estimated between 2,627 and 6,784 million lbs in 2006.
- Exploitation: The stock fishing mortality on the stock is estimated between 0.077 and 0.245 per year in 2006.

Atlantic stock:

- Stock abundance and biomass: The stock is estimated between 1,064 and 4,026 million lbs in 2006.
- Exploitation: The stock fishing mortality on the stock is estimated between 0.148 and 0.555 per year in 2006.

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

Methods used to calculate population benchmarks and management parameters followed guidelines provided by Restrepo *et al*. 1998, and proposed/alternative procedures described in Section I of the SEDAR 16 Stock Assessment Report. The VPA base cases for the Gulf of Mexico and Atlantic migratory stocks were proposed as the appropriate models to use for management advice.

The minimum spawning stock size threshold (MSST) was defined as [(1-M) or 0.5 whichever is greater]*B_{MSY}, the default for data-moderate situations in the guidelines. The maximum fishing mortality threshold (MFMT) was defined as F_{MSY} in the proposed/alternative procedures.

$F_{30\%SPR}$ was used as a proxy for F_{MSY}. $B_{30\%SPR}$ yield-per-recruit calculations require a spawner-recruit relationship, and this was estimated using VPA recruitment estimates, Beverton-Holt and an assumed steepness value of 0.95.

The current selectivity pattern used for yield per recruit calculations was derived from a normalization of the current F vector. Current F was calculated from the geometric mean of the age-specific F values from VPA for the most recent three years (2004-2006). The RP recommends that in the future reference F values be calculated by averaging across ages rather than using apical F. Average F values are easier to understand.

Yield per recruit calculations also require life history values for M, weight at age, maturity and fecundity. The values used for these were the same as for the VPA.

Within the VPA and also yield per recruit calculations, SSB was computed as numbers at age times maturity times fecundity to reflect egg production rather than biomass. The RP noted that yield per recruit calculations in particular are often made in relation to biomass rather than egg production, but agreed that incorporation of fecundity information is an improvement to the more usual procedure. The RP points out that fecundity needs to be well sampled before its use as a replacement for spawning biomass in an assessment will improve the calculation of reference points (See section 1.3.10).

Proposed alternative values of optimum yield were 65%, 75%, and 85% of F_{MSY}.

ABC values were provided using a range of constant F projections over the period 2007-2016. The constant F values were $F_{current}$, F_{MSY} (=$F_{30\%SPR}$), $F_{40\%SPR}$, 65% $F_{30\%SPR}$, 75% $F_{30\%SPR}$ and 85% $F_{30\%SPR}$. The assessment team provided this range as they believed that the selection of an appropriate ABC was a management decision.

The RP requested sensitivity tests of results using an alternative lower steepness value of 0.75.

Table 4.1 Management reference points from the Gulf of Mexico uncorrected base case using alternative steepness values

Steepness	$F_{30\%SPR}$	F_{MSY}^{*}	$B_{30\%SPR}$	B_{MSY}^{*}
0.95	0.25	0.41	2,941	1,709
0.75	0.25	0.23	2,393	2,676

* calculated using the Sissenwine and Shepherd (1987) approach that includes the SR relationship (not used in current management recommendations, but provided for comparison purposes).

Some reference point values are sensitive to the chosen steepness value. The RP had some concern that the recruitment estimates from VPA were uninformative about steepness, and that the default steepness value of 0.95 was arbitrarily chosen. The RP has made a research recommendation to improve the procedure for selecting an appropriate steepness value. In addition, the RP noted that a decrease in steepness produces a lower

MSST as this was based on the B_{MSY} proxy of $B_{30\%SPR}$. This counter-intuitive behavior is due to the B_{MSY} proxy being used, and possibly also fixing of the maximum expected recruitment level when fitting the stock recruitment relationship. The RP recommends that the behavior of the current control rules in relation to steepness be investigated using simulation as a research task, to test that the rules achieve management objectives as expected. Additionally, improved behavior at lower steepness values could be achieved by fitting the SR curve through an equilibrium point, rather than by limiting maximum recruitment.

$F_{40\%SPR}$ is considered to be an acceptable F_{OY} value in other US regions and other countries for the purpose of ABC calculations. The F_{OY} values of 65%, 75% and 85% of $F_{30\%SPR}$ represent different levels of conservativeness in the same range as $F_{40\%SPR}$. The use of different F_{OY} values appears to have been accepted in this fishery without investigation of the properties of each through simulation testing. The RP also recommends that operational methods to exploit the fishery at F_{OY} be tested by simulations.

Base case and plausible sensitivity results in Tables 2.3.1 and 2.3.2 showed that estimates of B_{2006}/MSST were robust, and indicated that both the GOM and ATL 2006 spawning stock biomass levels were above the MSST, and therefore not overfished. The range of plausible sensitivity results for the GOM for $F_{current}$ /MFMT was from 0.826 to 1.477 indicating that it is uncertain whether the stock is currently experiencing overfishing. For the ATL stock, $F_{current}$/MFMT plausible values were in the range of 0.615 to 1.085 indicating that if overfishing is occurring, it is at a low level. The run using the single fishery independent index was considered by the AW to be unreliable because of the limited data, and the resulting high estimate of $F_{current}$/MFMT ratio of 2.121 is not plausible. A range of possible ABC values are provided in these results, from 6.557 - 10.225 million lbs (excluding the Mexico landings) for the GOM and 7.610 - 9.850 million lbs for the ATL. The RP did not agree that the base case results provided central values within the plausible range of results examined, and have recommended that this uncertainty be incorporated into the TAC setting process.

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

Projection methods followed the recommendations of the AW, using a Beverton-Holt SR function based on VPA recruitment estimates, setting maximum expected recruitment to the geometric mean of the estimated recruits over the years that they were available (1981-2004 GOM and 1989-2004 ATL), and an assumed steepness value of 0.95. To estimate projection variance, 1000 bootstraps were run, using the CV of the observations to vary predicted recruitment about the fixed SR curve. Seven different projection scenarios were examined, using different levels of future catch: $F_{current}$, F_{MSY}=$F_{30\%SPR}$, $F_{40\%SPR}$, 65%$F_{30\%SPR}$, 75%$F_{30\%SPR}$ and 85%$F_{30\%SPR}$.

The RP agreed that the bootstrap procedure is adequate for estimating parametric uncertainty for the base model and catch scenario combinations. However, most of the

uncertainty in assessment outcomes is among alternative plausible model structures rather than within-model uncertainty.

Given this uncertainty across different model structures, the RP does not believe that error estimates from any single base model appropriately captures the uncertainty in the ABC and other stock condition indicators that result from this assessment.

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

The RP found that the uncertainty caused by observation error is adequately characterized. The empirical bootstrap approach is appropriate for these complex models. Uncertainty in catch estimation was not addressed in the VPA model, as it is assumed catch-at-age is known exactly. Index observation error was used as the basis for the bootstrap simulation. In SS3, other types of error can be addressed, including catchability process error. For the Atlantic stock, the RP considered that the estimated CVs of the indicators were unrealistically tight.

The current models, at least for the GOM, showed strong retrospective patterns. Retrospective patterns measure the ability of the stock assessment to forecast accurately and indicated that, in this case, there was a perceived change in productivity inconsistent over the time series. Retrospective patterns generally can be linked to some model misspecification. There are no simple solutions to this problem, and it adds to the level of uncertainty in the assessment.

Structural error needs to be addressed. It represents the difference between the model and reality, and is generally considered the most significant source of error. In this case, the RP has recommended developing a decision table with “states of nature” and the likely range of possible outcomes not exceeding the plausible range.

The RP recommended that the AW should present assessment results in the form of a decision table that represents the estimates derived from several plausible models that bracket the likely range of outcomes from the decision making. This, most importantly, considers the costs of making a decision assuming one hypothesis is true when an alternative hypothesis is closer to the real “state of nature”.

In the case of the GOM there were five models available. The alternative models (“Fishery Independent Indices Only” and “Fishery Dependent Indices Only”) were bracketed between the models featured in the table, and concern was expressed by members of the AW that the “Fishery Independent Indices Only” fit was poorly determined and unreliable.

The structure for the decision table is suggested as:

Gulf of Mexico Stock excluding Mexican catches

Decision (ABC/TAC) Mill. Pounds	"States of Nature"		
	Fishery Dependent Indices Only	MLE Indices Weighting	Corrected Base
8.305	P_{11}	P_{12}	P_{13}
9.396	P_{21}	P_{22}	P_{23}
10.335	P_{31}	P_{32}	P_{33}

In the case of the GOM stock, two model specifications where sensitive to initial specifications of terminal F (Fishery Independent Indices only and excluding the last 5 years of the SEAMAP groundfish survey, see Table 2.3.1) and two others were not comparable across management actions (Including Mexico catches and excluding Age 0 from the model), leaving the three possible models only.

Atlantic Stock

Decision (ABC/TAC) Mill. Pounds	"States of Nature"		
	MLE Indices Weighting	Base	New Index
8.331	P_{11}	P_{12}	P_{13}
8.662	P_{21}	P_{22}	P_{23}
9.850	P_{31}	P_{32}	P_{33}

In the tables above P_{ij} = probability that F_{2008} > MFMT which can be obtained from the bootstrap procedure for each model fit (to be completed by the AW), the decision is based on the ABC calculation (Yield equation 0.85*F30%SPR) on the allowable catch made by the council, which are produced from the models based on the formal decision rules. These are for guidance only on the final ABC recommendation from the SSC. The Base Model is considered the model closest to reality by the AW.

In future, the AW should consider developing alternative plausible hypotheses to their base case to aid the review process. The AW should consider and advise on the major uncertainties in the assessment. If a single dimension can be identified as the main source of uncertainty (e.g. steepness, productivity, weighting of abundance indices), this can be used to profile across this uncertainty for inclusion in a decision table.

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

The stock assessment and summary report is clearly and accurately presented. However, the RP recommended that the SEDAR develops procedures that minimize the burden on the various workshops to produce documentation. The number and length of documents

produced for review were extensive and complete, but clearly time consuming to produce and so extensive that it was difficult to identify the key areas which the assessment was sensitive to. An alternative approach is to develop a single document, adding, changing and editing sections as necessary so that the versions of the document represent a "snap-shot" of the current thinking. While a history of the assessment process is useful in theory, following this progression in detail is usually beyond any but a few involved intimately in the process. A good example of the updated document approach is the North Sea Herring Working Group Report (https://www.ices.dk/iceswork/wgdetailacfm.asp?wg=HAWG).

8. Evaluate the SEDAR Process. Identify any Terms of Reference which were inadequately addressed by the Data or Assessment Workshops; identify any additional information or assistance which will improve Review Workshops; suggest improvements or identify aspects requiring clarification.

The evaluation of the SEDAR workshops in addressing their terms of reference are in Tables 8.1 and 8.2. Overall, the workshops have conducted their work very conscientiously. They have clearly been professional and addressed almost all of the ToRs as well as might be expected. However, not all terms of reference were fully addressed.

The data workshop is required to "Evaluate the degree to which available indices adequately represent fishery and population conditions." (ToR 3) This was certainly done at a sampling / statistical level, but guidance was limited on how well these different indices reflect abundance.

The data workshop is required to "Provide maps of fishery effort and harvest." (ToR 4) These maps were not provided, although information on the spatial distribution of catch and effort was provided.

The assessment workshop ToR "Evaluate the results of past management actions and, if appropriate, probable impacts of current management actions with emphasis on determining progress toward stated management goals" was not met due to time constraints. However, the RP understand that the complexity of this task is very great and it is not feasible to be conducted in the time available.

Several data workshop ToRs (DW ToR 2, 3, 4) refer to "adequacy" of input information. The focus of the workshop was to provide the best information available, which is succeeded in doing. However, "adequacy" requires subjective judgment and is suitable for developing a base case assessment. What is also of interest to the assessment and review panels should be measures of uncertainty. Information helping identify the least reliable source of information among the catches, indices of abundance and size/age compositions or alternative inputs where "data" are estimated, might be used to develop alternative models to test for sensitivity. It should be noted that alternative models were suggested by the DW to test stock structure.

In the opinion of the RP, the AW TORs 6 and 8 contained inappropriate references to stock structure. Stock structure should be determined on scientific grounds, and is the

prerogative of the DW and AW, based on the scientific evidence and expert opinion only. Other mechanisms should exist for determining how these resources are shared among stakeholders.

The RP recommended that SEDAR attempts to evaluate the effectiveness of past management actions, as this provides feedback control important to this sort of process. The management actions have been listed, but there have not been evaluations except in the sense of the impact on monitoring indices. SEDAR should also develop standardized procedures to guide AW on methodology and especially on the presentation of results. This should include for example:

- Standard residual plots including QQ plots;
- Fish stock parameters presented in a standard way, e.g. arithmetic mean across ages as recommended here;
- Results of plausible alternative model fits in the form of a decision table

Table 8.1. Review of the Data Workshop terms of reference.

	Terms of Reference	Comments
1.	Characterize stock structure and develop a unit stock definition. Provide maps of species and stock distribution.	**Met**: The available evidence on stock structure is provided and a hypothesis on population structure, including a map, is proposed.
2.	Tabulate available life history information (e.g., age, growth, natural mortality, reproductive characteristics); provide appropriate models to describe growth, maturation, and fecundity by age, sex, or length as applicable. Evaluate the adequacy of available life-history information for conducting stock assessments and recommend life history information for use in population modeling.	**Met**: Life history information is provided and is complete. Information is provided in a form suitable for stock assessment. The data workshop provided clear recommendations on which information to use. Overall uncertainty among information was not characterized (DW Final Section 2.11).
3.	Provide measures of population abundance that are appropriate for stock assessment. Document all programs used to develop indices, addressing program objectives, methods, coverage, sampling intensity, and other relevant characteristics. Provide maps of survey coverage. Consider relevant fishery dependent and independent data sources; develop values by appropriate strata (e.g., age, size, area, and fishery); provide measures of precision. Evaluate the degree to which available indices adequately represent fishery and population conditions. Recommend which data sources should be	**Met except the evaluation of indices was limited**: Measures of abundance were provided, reviewed and recommendations were made on which indices were appropriate for the stock assessment. Methods used to estimate indices were documented providing appropriate area, age groups, and sampling precision relevant to indicate how the population models should reference indices. Indices were not evaluated with respect to their relationship to and ability to track overall abundance.

	considered in assessment modeling.	
4.	Characterize commercial and recreational catch, including both landings and discard removals, in weight and number. Evaluate the adequacy of available data for accurately characterizing harvest and discard by species and fishery sector. Provide length and age distributions if feasible. Provide maps of fishery effort and harvest.	**Met**: Commercial and recreational catches are well described, including both discards and landings, and the adequacy of the data has been evaluated. Age, sex and length distributions have been provided where possible.
5.	Provide recommendations for future research in areas such as sampling, fishery monitoring, and stock assessment. Include specific guidance on sampling intensity and coverage where possible. Provide discussion of progress on research and monitoring recommended by SEDAR 5.	**Met**: Extensive research recommendations were provided covering all relevant areas.
6.	Prepare complete documentation of workshop actions and decisions (Section II. of the SEDAR assessment report).	**Met**: A complete document was prepared and given to the SEDAR assessment group.

Table 8.2. Review of the Assessment Workshop terms of reference.

	Terms of Reference	Comments
1.	Review any changes in data following the data workshop and any analyses suggested by the data workshop. Summarize data as used in each assessment model. Provide justification for any deviations from Data Workshop recommendations.	**Met**: The AW reviewed changes and recommendations from the DW. Where changes have been made, these were documented and explained.
2.	Develop population assessment models that are compatible with available data and recommend which model and configuration is deemed most reliable or useful for providing advice. Document all input data, assumptions, and equations.	**Met**: The VPA implemented in VPA2Box was selected. The SS3 model which has been developed could not be completed in time, although it was considered to be more realistic description of population processes. The model has been fully documented.
3.	Provide estimates of stock population parameters (fishing mortality, abundance, biomass, selectivity, stock-recruitment relationship, etc); include appropriate and representative measures of precision for parameter estimates.	**Met**: Estimates of stock population parameters with bootstrap estimates of precision have been provided.
4.	Characterize uncertainty in the assessment and estimated values, considering components such as input data, modeling approach, and model configuration. Provide appropriate measures of model performance, reliability, and 'goodness of fit'.	**Met**: The uncertainty of the different runs is assessed using a bootstrap as well as reporting standard fit diagnostics.
5.	Provide yield-per-recruit, spawner-per-recruit,	**Met**: YPR, SPR were provided. SPR

	and stock-recruitment evaluations.	calculations provided the reference point. A stock-recruitment model was proposed and used in the projections.
6.	Provide estimates for SFA criteria consistent with applicable FMPs, management programs, and National Standards. This may include: evaluating existing SFA benchmarks, estimating alternative SFA benchmarks; and recommending proxy values. SFA parameters shall be provided for the Gulf and Atlantic Migratory Units as currently defined using the most current mixing data.	**Met**: SFA benchmarks were reviewed and calculated for the relevant stocks.
7.	Provide declarations of stock status relative to SFA benchmarks.	**Met**: The stocks' status was evaluated with respect to the reference points.
8.	Estimate Allowable Biological Catch (ABC) based on the following criteria: A) Based on migratory groups and mixing zone dynamics defined using best available scientific information, provide separate ABC values for each of two management areas delineated at the Miami-Dade/Monroe County line: all fish caught north of the line allocated to the Atlantic management area and all fish caught south of the line allocated to the Gulf management area. B) Based on migratory groups and mixing zone dynamics as currently defined, provide separate ABC values for the Gulf and Atlantic Migratory Units based on allocating all fish in the mixing zone to the Gulf Migratory Unit (essentially the 'continuity' approach). C) Based on migratory groups and mixing zone dynamics as currently defined, provide separate ABC values for the Gulf and Atlantic migratory units based on allocating 50% Gulf of Mexico and South Atlantic King Mackerel of the fish in the mixing zone to the Gulf Migratory Unit and 50% of the fish to the Atlantic Migratory Unit. D) Based on migratory groups and mixing zone dynamics defined using best available scientific information, provide separate ABC values for each of two management areas delineated at the Gulf and South Atlantic Council boundaries	**Met**: ABCs were calculated for the 4 stock mixing scenariosmanagement boundary scenarios.

9.	Project future stock conditions (biomass, abundance, and exploitation) and develop rebuilding schedules if warranted; include estimated generation time. Stock projections shall be developed in accordance with the following: A) If stock is overfished: F=0, F=current, F=Fmsy, Ftarget (OY), F=Frebuild (max that rebuild in allowed time) B) If stock is overfishing F=Fcurrent, F=Fmsy, F= Ftarget (OY) C) If stock is neither overfished nor overfishing F=Fcurrent, F=Fmsy, F=Ftarget (OY)	**Met**: Projections have been carried out as required. No rebuilding is warranted.
10.	Evaluate the results of past management actions and, if appropriate, probable impacts of current management actions with emphasis on determining progress toward stated management goals.	**Not met**: The time constraints prevented the assessment panel from evaluating past management. The RP felt that this term or reference was not realistic.
11.	Provide recommendations for future research and data collection (field and assessment); be as specific as practicable in describing sampling design and sampling intensity. Provide discussion of progress on research and monitoring recommended by SEDAR 5.	**Met**: Extensive recommendations have been provided for future research. The main SEDAR 5 research recommendation, moving to a statistical catch-at-age model, was addressed.
12.	Complete the Assessment Workshop Report (Section III of the SEDAR Stock Assessment Report) and prepare a first draft of the Summary Report.	**Met**: The assessment workshop report was completed.

9. Review the research recommendations provided by the Data and Assessment workshops and make any additional recommendations warranted. Clearly indicate the research and monitoring needs that may appreciably improve the reliability of future assessments. Recommend an appropriate interval for the next assessment.

The assessment and data workshops have identified the most important research required to improve the assessment. Those areas of research requiring highest priority as well as some additional research are outlined below, based on the need to appreciably improve the reliability of future assessments. Where possible, this research should be completed for the next assessment.

The RW emphasized the importance of the Mexican catches. This was addressed by the AW's recommended research, to determine whether separate stocks exist in the eastern and western portions of the GOM and the relationship of king mackerel off the coast of

Mexico with U.S. king mackerel stocks (DW 2 & 3; AW 3, 4 & 5). The RW considered these a priority.

An objective procedure to justify the choice of steepness value used for king mackerel modeling is required. This may be either from best fits to available data, or choice of appropriate values for similar species from a meta-analysis. It should also be investigated whether improved behavior at lower steepness values could be achieved by fitting the SR curve through an equilibrium point, rather than by limiting maximum recruitment. This applies both to reference point calculation and projections.

The RW was concerned with the accuracy of the available abundance indices. With the exception of the research to remove the suspected bias in the log-book data (AW 8), no recommendations on improving the abundance indices were made by either the DW or AW. Given the problems with the indices, research should include identifying methods which might improve collection and standardization of data used for this purpose. In particular, the RW believed that improved stock-wide fishery independent indices may be required to carry out control to the level of precision implied by management. It is also important that the commercial logbook index constructed for the Atlantic stock unit is used if possible in future assessments.

The RP recommended that the behavior of the current control rules that use per recruit $F_{30\%SPR}$ values be investigated using simulation, to ensure that they achieve management objectives as expected. A useful framework for this form of testing is known as management strategy evaluation that includes an operating model of fish population dynamics (using various plausible scenarios), fisheries scientific sampling from the population with error, fishing fleet operations and catch, stock assessment and management action as simulation components (e.g. see ICES Marine Science Symposia, 1999).

The RP endorses the AW recommendation that the discrepancy between the two programming codes R and SAS that were used in SEDAR5 and SEDAR16, respectively for estimating shrimp trawl bycatch be resolved.

If the development of the SS3 model is to continue, research programs are required that improve monitoring of the stock mixing. These include tagging studies, otolith microchemistry and shape analysis studies, and the collection of microsatellite genetic marker data to determine mixing rates (DW 1; AW 6 & 7).

Otoliths from the mixing zone need to be evaluated with shape or elemental analyses in order to assign them to one of the two stocks for use in future assessments.

The size and age maturity functions should be updated as the most recent estimates are over 20 years old.

Either the intensity of sampling for fecundity should be greatly increased, or else weight-at-age of mature fish should be used as a proxy for spawning potential.

Procedures should be investigated for incorporating uncertainty and assign utility across model structures into ABC and stock condition calculations. Most of the uncertainty in assessment outcomes is between alternative plausible model structures.

An important uncertainty for the GOM stock is whether a series of recent good recruitments that appear in some indices will contribute in the medium term to increase stock biomass of fish of a size targeted by the commercial and recreational sectors. It will take two to three years for these fish to enter the fishery and for a stock assessment to determine what the impact of those recruitments really is. Therefore, the RP recommends that an update assessment be conducted in two to three years.

The SEDAR Steering Committee should investigate the methodology currently used by the National Hurricane Center to develop consensus forecast models from varied different forecast models to determine if a similar approach is suitable for in improving estimates of stock status and medium term management forecasts with more realistic estimates of uncertainty than can be gained from an examination of internal variability within a single model.

References

Finucane, J.H., L. A. Collins, H.A. Brusher, and C.H. Saloman. 1986. Reproductive biology of king mackerel, Scomberomorus cavalla, from the southeastern United States. Fishery Bulletin 84(4):841-850.

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

ICES Marine Science Symposia. 1999. *Confronting Uncertainty in the Evaluation and Implementation of Fisheries-Management Systems*. (Cape Town, 1998). ICES Journal of Marine Science, vol. 56, no. 6.

Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology 49: 627-647.

Restrepo, V. R., G. G. Thompson, P. M. Mace, W. L. Gabriel, L. L. Low, A. D. MacCall, R. D. Methot, J. E. Powers, B. L. Taylor, P. R. Wade, and J. F. Witzig. 1998. Technical Guidance on the Use of Precautionary Approaches to Implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act. NOAA Tech. Memo. NMFS, Scientific Publications Office (F/SPO) 31:1-54.

Sissenwine, R.P. and Shepherd, J.G. 1987. An alternative perspective on recruitment overfishing and biological reference points. Can. J. Fish. Aquat. Sci. 44:913-918.

3. APPOINTEE STATEMENTS

(Any written comment or opinion statements submitted by appointed participants or observers)

Comments on the SEDAR 16 King Mackerel Review Workshop, August 4-8, 2008, Jacksonville, Florida

I attended the Review Workshop as an Observer from the Gulf Council SSC. My impressions of the meeting are:

1. The Review Panel Members were thoroughly knowledgeable in the techniques and tools of modern population biology. They were knowledgeable about the reports from the SEDAR 16 King Mackerel Data and Assessment Workshops, which formed the basis for discussion at the meeting.

2. The Assessment Team from the Southeast Fisheries Science Center (and one member from the Alaska Fisheries Science Center) presented the results of the Assessment Workshop in a clear and concise manner. The Team responded to requests from the Review Panel for new model runs to test the robustness of the Assessment Workshop results. These new model runs were usually available and presented to the Review Panel within a few hours.

3. The questions from the Review Panel were penetrating, but were responded to well by the Assessment Team. Both groups recognized the limitations of the data and the difficulties of reaching definitive conclusions about the status of a stock of migratory fish ranging over a wide area and subject to multiple fisheries in different areas.

4. Conflicting data trends and data gaps presented the Panel with difficult decisions. In particular, the Panel noted the conflicting trends between fishery independent data (generally increasing trends) and fishery dependent data (generally level or slightly decreasing trends). Also, the Panel noted data gaps such as the lack of catch information from the Mexican fisheries, and the lack of coverage in the eastern Gulf of Mexico of the SEAMAP Groundfish Survey (GOM) (used as an index of Gulf Group age 0 king mackerel) and considered the low nominal numbers of larvae collected in the Fall Plankton Survey (GOM) (used as an index of spawning stock biomass for the Gulf Group). Nevertheless, there was a great deal of information to consider and use in the analyses. I consider the data and analysis of SEDAR 16 an improvement over the previous assessment of king mackerel.

5. The meeting proceeded in an orderly manner. Observers listened intently during the presentations and during discussions between the Panel and Assessment Team. Observers contributed when recognized by the Chair and when appropriate. In particular, the contributions of industry observers Ben Hartig and Bob Zales were informative and helpful. During brief intermissions, such as when the Group was waiting for the results of new model runs, Observers quietly and without disturbance to the Group reviewed materials on their computers and Panel Members either reviewed materials or discussed next steps in the meeting.

6. Meeting materials (Presentations and background materials) were available on the internal network and the network functioned without problems during the meeting.

7. At the conclusion of the meeting, the Review Panel presented a short, preliminary summary of their findings.

Albert Jones
Member, Scientific and Statistical Committee (SSC)
Gulf of Mexico Fishery Management Council
August 18, 2008