SEDAR

Southeast Data, Assessment, and Review

SEDAR 21

HMS Atlantic Blacknose Shark

SECTION V: Review Workshop Report

May 2011

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

Table of Contents

Table of Contents 2
1. INTRODUCTION 2
1.1 WORKSHOP TIME AND PLACE 2
1.2 TERMS OF REFERENCE 2
1.3 LIST OF PARTICIPANTS 3
2. REVIEW PANEL REPORT 4

1. INTRODUCTION

1.1 WORKSHOP TIME AND PLACE

The SEDAR 21 Review Workshop was held April 18-22, 2011 in Annapolis, MD.

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 stock status(*e.g., MSY, Fmsy, Bmsy, MSST, MFMT, or their proxies*); recommend appropriate management benchmarks, provide estimated values for management benchmarks, and declare stock status, consistent with the stock status determination criteria, benchmark, and biological reference points in the Consolidated HMS FMP, proposed FMPs and Amendments, other ongoing or proposed management programs, and National Standards.
5. Evaluate the adequacy, appropriateness, and application of the methods used to project future population status, rebuilding timeframe, and generation time; recommend appropriate estimates of future stock condition (e.g., exploitation, abundance, biomass).
6. Evaluate the adequacy, appropriateness, and application of methods used to characterize the 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 and that reported results are consistent with Review Panel recommendations. If there are differences between the AW and RW due to reviewer's requests for changes and/or additional model runs, etc., describe those reasons and results.

8. Evaluate the SEDAR Process as applied to the reviewed assessments and identify any Terms of Reference that were inadequately addressed by the Data or Assessment Workshops.
9. Consider the research recommendations provided by the Data and Assessment workshops and make any additional recommendations or prioritizations warranted. Clearly denote research and monitoring needs that could improve the reliability of future assessments. Recommend an appropriate interval for the next assessment, and whether a benchmark or update assessment is warranted.
10. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Provide a list of tasks that were not completed, who is responsible for completing each task, and when each task will be completed. Complete and submit the Final Summary 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

Workshop Panel

Larry Massey, Chair.....NMFS SEFSC
Jamie Gibson.....CIE Reviewer
Neil Klaer.....CIE Reviewer
Shelton Harley .....CIE Reviewer

Analytic Representation

Enric Cortés ..... NMFS SEFSC Panama City
Kate Andrews.....NMFS SEFSC Beaufort
Paul Conn.....NMFS AFSC

Rapporteur

Ivy Baremore ..... NMFS SEFSC Panama City

HMS Representation

Karyl Brewster-Geisz.....NMFS HMS

Observers

Peter Cooper.....NMFS HMS
Chris Vonderweidt.....ASMFC

Staff

Julie Neer ..... SEDAR
Tyree Davis.....NMFS Miami

2. REVIEW PANEL REPORT

Executive Summary

An independent peer-review panel workshop (RW) convened in Annapolis, Maryland, April 18-22, 2011, in a public meeting to review draft stock assessments for four SEDAR 21 stocks of Highly Migratory Species: U.S. South Atlantic blacknose, U.S. Gulf of Mexico (GoM) blacknose, sandbar, and dusky sharks. Data and assessment reports were presented that had been developed since June 2010 in a data workshop (DW) and a series of assessment webinars (AW). After examining the documents, the review panel (RP) questioned and asked for additional information from the chief stock-assessment analysts to determine whether the decisions, assumptions, assessment models, and outputs were adequate, complete, and scientifically sound. The RP's findings are:

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

Data: The RP examined all input parameters and data used in the four assessments. Uncertainties associated with some of the sources of information were addressed by the assessment team (AT) via sensitivity analysis.

Stock Units: The AT used genetic studies, life history characteristics and tagging information to develop one stock unit for sandbar, one for dusky, and two for blacknose (GoM stock and Atlantic stock). The RP accepted these, but noted that if the sampling for the genetic and tagging studies occurred at a time when discrete populations were mixed, finer scale stock structure would not be detected. Additionally, if low rates of straying exist between populations or if genetic divergence is recent, genetic studies may not detect population structure even if populations are demographically uncoupled.

Life History Parameters: The RP examined the biological characteristics for the four stocks and generally accepted that the information was used appropriately, but expressed reservations about how the limited age and length data were used to derive selectivities for the various fisheries and surveys. Some selectivity-related sensitivity analyses were undertaken.

Abundance Indices: The RP generally accepted the selection of indices, agreed with the AT that many of the indices were short relative to the life span of the stocks under assessment, that most exhibited annual variability that exceeds what might be expected for these stocks, and that several did not span the full geographic distribution of the stocks. Assessment results were sensitive to these changes, and the RP appreciated that these variations were carried forward as sensitivity analyses.

Landings and Removals: Data issues identified by the data workshop panel include: under-reporting, species identification, spatial coverage, landings being aggregated for more than one species and whether data were included in more than one database creating the potential for double counting. The AW evaluated the effects of under- or overestimating landings and removals with model runs using higher and lower landings. The RP agreed that this approach was a reasonable way to evaluate how model output is scaled to overall abundance, but noted that the approach would only work if over- or under-reporting, or other issues with landings and removals data were similar over the entire time series. Regarding the historical period—defined as the period from a year in which the population could be considered to be at virgin levels, to the time at which landings data become available—the AT assumed that fishing effort increased during this period, and explored the effects of the assumed nature of this increase (e.g. linear versus exponential) on the assessment results using additional model runs with different assumptions. The RP accepted this as a reasonable approach, but also agreed with the AT that there was considerable uncertainty about the removals during the historical period and therefore the status of the populations at the time when landings data became available.

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

Assessment Models: Two assessment methods were used across the four stock assessments, both being variants on the basic Age-Structured Production Model. For sandbar and the two blacknose stocks, a state-space variant was used: the state-space age-structured production model (referred to here as ASPM); while for the dusky assessment the catch-free variant was used: Age-Structured Catch-Free Production Model (ASCFM).

Overall it is the conclusion of the RP that these approaches are appropriate to the assessment of these stocks given some of the particular issues that relate to shark assessments, e.g., uncertain catch histories, and low productivity.

ASPM – Sandbar and Blacknose Assessments: The RP concludes that the general population dynamics assumed in the assessment model are appropriate for the assessments, but it is not clear that the added complexity of the state-space age-structured production model variant of ASPM was useful. The RP believes that using a model that did not include added complexity would not change the conclusions of the assessments or this review but would have made the models and model results easier to interpret, in particular allowing for easier analysis of the impacts of alternative data weighting scenarios.

ASCFM – Dusky Shark: The underlying population dynamics for ASCFM are essentially the same as those for ASPM – the major difference in the approaches being that absolute estimates fishery removals are not used within ASCFM. This is the same modeling platform used for the previous dusky assessment and is necessary due to the difficulties in obtaining any sort of catch history. A concern of the RP was the lack of simulation studies to determine the performance of the ASCFM under known conditions, particularly the response of the model to alternative scenarios about productivity when catches are known. This is further discussed under ToR 9.

GoM Blacknose Shark: Sensitivity runs requested by the RP demonstrated that the blacknose GoM assessment model was unable to fit apparent trends in the abundance indices at all, unless implausible additional historical catches were also estimated. This fundamental lack of fit of the model to the input data caused the RP to reject the blacknose GoM assessment model.

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

The RP accepted the model results for Atlantic blacknose, sandbar and dusky sharks although there was considerable uncertainty in those results that is conveyed as part of the RW recommendations.

ToR 4: Evaluate the methods used to estimate population benchmarks and stock status (e.g., MSY, FMSY, BMSY, MSST, MFMT, or their proxies); recommend appropriate management benchmarks, provide estimated values for management benchmarks, and declare stock status, consistent with the stock status determination criteria, benchmark, and biological reference points in the Consolidated HMS FMP, proposed FMPs and Amendments, other ongoing or proposed management programs, and National Standards.

Atlantic Blacknose Shark: Results showed that the stock was overfished (SSF_{2009}/SSF_{MSY} of 0.43 to 0.64, all below MSST) and therefore subject to rebuilding. Current F values over all sensitivities also indicated that the stock was subject to overfishing (F_{2009}/F_{MSY} of 3.26 to 22.53).

Sandbar Shark: Results showed that the stock was overfished and therefore subject to rebuilding. Current F values over most sensitivities indicated that the stock was not currently subject to overfishing (F_{2009}/F_{MSY} 0.29 to 0.93). However, the low productivity scenario did indicate overfishing (F_{2009}/F_{MSY} of 2.62).

Dusky Shark: Results showed that the stock was overfished (SSB_{2009}/SSB_{MSY} of 0.41 to 0.50) and therefore subject to rebuilding. Current F values over all sensitivities also indicated that the stock was subject to overfishing (F_{2009}/F_{MSY} of 1.39 to 4.35).

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

Given the critical importance of incorporating uncertainty in the projections, the RP did not accept the projections or projection method applied to the sandbar and Atlantic blacknose stocks. However, it is the opinion of the RP that the methodology applied to the dusky assessment is sufficient and those projection results valid.

ToR 6: Evaluate the adequacy, appropriateness, and application of methods used to characterize the 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 concluded that the AT has used and applied appropriate methods to characterize uncertainty in the four stock assessments.

Future assessments could consider additional approaches to characterize uncertainty.

The approaches used by the AT are appropriate and the RP has used the information provided by the AT to characterize the bounds of uncertainty in current and projected stock status under ToRs 3-5.

ToR 7: Ensure that stock assessment results are clearly and accurately presented in the Stock Assessment Report and that reported results are consistent with RP recommendations. If there are differences between the assessment workshop and review workshop due to reviewer's requests for changes and/or additional model runs, etc., describe those reasons and results.

Assessment documentation prepared by the AT was comprehensive for use of input data, model description, results and sensitivity analysis selection.

Considerable improvement is needed in the provision of model diagnostics. Evidence of convergence should be included particularly for the base case, minimally in the form of convergence statistics and preferably as MCMC diagnostic plots. Sensitivity results should include for the base case and all sensitivities, as columns, the individual objective function components. Developing national standards in stock assessment documentation should be implemented.

ToR 8. Evaluate the SEDAR Process as applied to the reviewed assessments and identify any Terms of Reference that were inadequately addressed by the data or assessment workshops.

The RP believes this SEDAR process has, overall, led to a comprehensive assessment of these stocks.

An independent reviewer participated in the data workshop and the AW report was also reviewed prior to being finalized for the RW. It remained unclear to the RP whether placing greater emphasis on reviews earlier in assessment processes will automatically lead to better assessments.

The RP believes that the ToRs of the data and AW were generally met.

There is a recommendation in the AW report that more time should be available for the data vetting process, while at the RW, the time available for the assessment modeling appeared to be one of the factors limiting further model development.

The review of four stocks in four days at the RW was only possible because three of the stocks used the same model and limited time was placed on the review of the data inputs.

The RP endorses the AT recommendation that no more than probably two stocks be assessed at one time with the same number of participants. The RP notes that the time required for a RW depends not only on the number of stocks, but also on the complexity of the individual assessments.

ToR 9. Consider the research recommendations provided by the data and assessment workshops and make any additional recommendations or prioritizations warranted. Clearly denote research and monitoring needs that could improve the reliability of future assessments. Recommend an appropriate interval for the next assessment, and whether a benchmark or update assessment is warranted.

The RP considers the following to be priorities:

- Research on post-release survival by fishing sector and gear type.
- Research on fecundity and reproductive frequency.
- Regular collection of age-specific data.
- Tagging studies conducted in collaboration with international entities (e.g. Mexico).
- Evaluation of the individual indices of abundance via power analyses.
- A small study on how to make the best use of the knowledge of the data workshop participants for developing index rankings.
- Ensuring information about sex, length and age is collected.
- Research that improves the understanding of historical landings, both in the modern and historical period and to support the assumptions about when stocks are at virgin biomass if this assumption is carried forward in future assessments.
- Improved observer coverage particularly during periods of regulatory or gear changes (e.g. TEDs).
- Further model development using simpler and more complex models.
- Estimating fishery and survey selectivities within the assessment model.
- Development of a two sex model.
- Fitting the model to either length or age data.
- Exploration of models that do not require an assumption that the population is at virgin levels at some point in time.
- If external age-length keys are used in future assessments, development of a key based on a growth model to better assign proportions-at-age in each length class.
- Simulation tests (management strategy evaluation) to test the performance of alternative assessment methods (including the catch-free model, ASPM, ASPIC, SS3, or stock specific models), recruitment parameterizations, harvest control rules, assessment frequency and data collection.
- For GoM blacknose shark, the appropriate interval of the next assessment depends on progress made towards reconciling the issues raised during this assessment process. For Atlantic blacknose shark, dusky shark and sandbar shark, the RP recognizes that population growth is expected to be relatively slow, but that modifications to the model may result in a different assessment of status. Benchmark assessments are recommended once the

modifications are made. Additionally, for dusky shark, given the retrospective patterns in the present analysis and the resulting uncertainties in the assessment, updates using the existing model in the shorter term are also recommended. In the longer-term, development of a set of indicators (age-structure, total mortality estimates from catch curves, changes in abundance indices values) that could be used to determine whether status has changed sufficiently to warrant a full assessment, is recommended.

ToR 10. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Provide a list of tasks that were not completed, who is responsible for completing each task, and when each task will be completed. Complete and submit the Final Summary Report within 3 weeks of workshop conclusion.

This report is the peer-review panel's summary of its evaluation of the stock assessments based on the given terms of reference.

SEDAR 21 Stock Assessment Review

A panel of three independent peer reviewers met in a workshop open to the public in Annapolis, Maryland, April 18-22, to critically examine population assessments for four shark stocks of Highly Migratory Species (HMS) (see list of participants in Appendix I). The workshop (SEDAR 21) was hosted by the office of the Southeast Data Assessment and Review (SEDAR)—an independent, cooperative fishery management council process for the southeastern United States. SEDAR was instituted in 2002 to improve the quality and reliability of fishery stock assessments in the U.S. South Atlantic, Gulf of Mexico (GoM), and U.S. Caribbean Sea by using a transparent and participatory process that typically includes federal, state, university and non-governmental scientists.

The stocks assessed in SEDAR 21 were dusky shark, sandbar shark, U.S. South Atlantic blacknose shark, and U.S. GoM blacknose shark. Starting in June 2010, scientists met in a data workshop and in a series of webinars to conduct a rigorous process of gathering and collating associated data, analyzing pertinent data and information, and conducting population model building and testing. Those efforts produced a stock assessment report for each fishery stock.

The objective of the independent review panel (RP) was to critically examine in a workshop the data, assumptions, decisions, analyses, models, and outputs for each stock. At the review workshop (RW), the panel studied the assessment documents and interviewed the chief stock assessment scientists to revisit each aspect of the assessment. The panel also requested and received additional sensitivity model runs and outputs.

The RP followed ten Terms of Reference (ToR). Since the four stocks were similar in nature, the same modeling software was applied to data and information for each stock, and the assessments were somewhat parallel in structure, the panel decided to provide one summary document that provides their consensus set of findings for all four stock assessments. Those findings are listed below.

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

The RP examined all input parameters and data used in these four assessments and commends the assessment team (AT) for providing thorough reviews of the multiple data sources available for these shark stocks. Data included abundance indices based on both fisheriey-dependent as well as fishery-independent surveys, commercial landings and removals including discards and by-catch, recreational fisheries landings and removals including discard mortality, length and age data, life history information such as age-at-maturation and fecundity, as well as the results of genetic studies and tagging information used to determine stock structure. As described below, there are uncertainties associated with some of these sources of information which the AT carried through the assessment via sensitivity analyses.

Stock Units

The AT used genetic studies, life history characteristics and tagging information, to develop the stock units used in these assessments. As appropriately described by the AT, there is relatively little information for delineating population structure within each of the assessed species. For sandbar and dusky sharks the AT determined that one stock would be appropriate because genetic data indicate no significant differentiation between the GoM and U.S. South Atlantic tag-recapture data showed a high frequency of movement between basins. For blacknose sharks, the AT determined two stocks would be appropriate because tagging studies showed very little movement between the GoM and the Atlantic Ocean, and because the life history information indicated that the reproductive cycle differed between these regions. The RP accepted these determinations, but noted that if the sampling for the genetic and tagging studies occurred at a time when discrete populations were mixed, finer scale stock structure would not be detected. Additionally, if low rates of straying exist between populations or if genetic divergence is recent, genetic studies may not detect population structure even if populations are demographically

uncoupled. A potential consequence of defining a larger stock is that exploitation at levels appropriate for the overall stocks could lead to overfishing of lower-productivity populations if they exist.

Life History Parameters

Although many aspects of shark life history are not well studied, others such as fecundity for some species are better known and can be informative in assessment models. The AT provided information on natural mortality, discard mortality, age and growth, as well as reproduction (maturity schedules, fecundity and reproductive cycles) for each of the four stocks. The RP examined the biological characteristics for the four stocks and generally accepted that the information was used appropriately. In particular, the RP appreciated that the uncertainties in the biological data was carried through the assessment via sensitivity analyses to evaluate how assumptions about natural mortality and reproductive frequency affected the assessment results. A discard mortality rate of 6% was used in the assessment. It was suggested during the review that as an alternative to a constant value, this value could be scaled up or down as a proportion of those landed dead, the idea being that the greater this proportion, the greater the stress on the animals landed alive and the lower the probability that they would live after being released.

Although the RP accepted the biological inputs as used in the assessment, the RP did express reservations about how the limited age and length data were used to derive selectivities for the various fisheries and surveys, an analysis that was carried out externally to the assessment model. Exploration at the RW showed that length measurements were not always the same (e.g. pre-caudal, total) requiring some standardization before it could be used. Because relatively small amounts of length data were available, data were at times aggregated over years prior to being converted to ages. Although this approach may be more appropriate for some sharks than some teleost species due to their lower reproductive rates, it does require the assumption that size-at-age and numbers-at-age have not changed during these years. Numbers-at-age particularly, could change rapidly if fishing mortality rates increase during the period over which these data are aggregated. The length frequencies were then converted to age frequencies using an age-length key. This key was developed using proportions at age in each length category. Because age data are sparse, this approach led to some anomalies in age-length keys. As an example, for dusky shark, sharks in the 180-190 cm category were assigned to age classes 5, 7, 8 and 10 with 25% of the sharks in each age class. The RP believed that the use of a growth model to assign an age distribution to sharks in a given length category would have led to a distribution that was more plausible biologically. Selectivity curves were then derived for the various fisheries and surveys using either algorithm or in many cases by fitting the selectivity by eye. This approach requires estimating the selectivity with knowledge of the true underlying age structure. An assumption is made that the fully selected age class is the one that is most abundant in the sample, an assumption that may not be valid if total mortality is high. The RP expressed a strong preference for including the length data and a growth model or age-length key in the assessment model in order that both abundance-at-age and selectivity can be estimated simultaneously overcoming this problem. However, the RP acknowledged that the only limited length data were available and that the assessment models being used were not set up to fit to these data. The AT did correctly acknowledge the subjectivities introduced by the method that was used to derive these curves. The RP accepted that, given the limited data and the models being used, the selectivity curves were sufficient for this assessment.

Abundance Indices

The AT considered a total of fifty-eight indices of abundance, based on both fishery dependent and fishery independent data, at the data workshop. Factors considered when selecting the indices to be used in the assessments included sample size, the length of the time series and whether the survey was conducted in appropriate habitat. The indices that were used are summarized in Table 1. The RP generally accepted the selection of indices, and agreed with the AT that many of the indices were short relative to the life span of the stocks under assessment, that most exhibited annual variability that exceeds what might be expected for these stocks and that several did not span the full geographic distribution of the

stocks. The AT also correctly pointed out that some indices had different trends, indicating, together with their variability, that they were likely responding to factors other than stock-level abundance. Additionally, the AT ranked the indices based on factors such as geographic coverage as a way of giving different weightings to the indices in the assessment model. The RP accepted the rankings, believing the data workshop participants understood best the indices and their strengths and weaknesses.

Table 1. Abundance indices fitted by the shark assessment models with rankings assigned by the data workshop.

Abundance Index	**Blacknose GoM**	**Blacknose Atlantic**	**Dusky**	**Sandbar**
Fishery independent				
Panama city gillnet	3(2)			4
VIMS longline			3	2
Dauphin Is sea lab BLL	5			
NMFS coastspan LL				2(3)
NMFS NE LL			1	2
GA coastspan LL/GADNR		4		4
SC coastspan/SCDNR		2		3(2)
NMFS historical LL			1s	
UNC LL		4	1s	
MOTE ML LL	3			
NMFS SE BLL	1	1		1
NMFS SEAMAP trawl	2(2)			
Fishery dependent				
SEFC shark BLL OB	4	3	1	2
Drift gillnet OB		3		
Sink gillnet OB		1s		
SE pelagic LL OB/log			2	2
MRFSS				
NE gillnet OB				
Coastal fishery gillnet log		4		
Coastal fishery BLL log				
Large pelagic survey			4	5

Note: values shown in brackets indicate multiple indices from the same source (e.g. adult/juvenile), s indicates that the index was used for a sensitivity run only.

The AT examined the influence of the indices on the assessment results in several ways: by a variety of weighting schemes such as weighting all data points equally, weighting individual data by the inverse of their CV, and weighting individual indices by their rank, by fitting the model using subsets of the indices (e.g. fishery independent indices only) and by deriving a single index using a hierarchical model. Assessment results were sensitive to these changes, and the RP appreciated that these variations were carried forward as sensitivity analyses.

Landings and Removals

Landings and removals for sandbar, dusky and blacknose sharks, including commercial landings, recreational landings, discards and discard mortality and bycatch, are difficult to estimate. Issues were well described in the data workshop reports, including: under-reporting, species identification, spatial

coverage, landings being aggregated for more than one species and whether data were included in more than one database creating the potential for double counting.

For sandbar shark, catches were included in the model as: the commercial and unreported catch series split into the GoM and Atlantic components, a recreational and Mexican catch series, and menhaden fishery discards. In the case of Atlantic blacknose shark, catches were included in the model as: commercial landings (bottom longlines, nets and lines) recreational catches, shrimp bycatch and bottom longline discards. For GoM blacknose shark, catches were included in the model as: commercial landings (bottom longlines, nets and lines) recreational catches, shrimp bycatch and bottom longline discards. Of these data, the shrimp bycatch in the GoM, which comprises most of the catches for this stock, is a key source of uncertainty particularly during and before the historical period defined for the model. For dusky shark, catches were considered to be too uncertain to be useful, leading to the use of a catch-free model. Instead, relative effort series were developed for the directed bottom longline, pelagic longline fishery and the recreational fishery.

The AW evaluated the effect of under- or over-estimating landings and removals using model runs with both higher and lower landings. The RP agreed that this approach was a reasonable way to evaluate how model output is scaled to overall abundance, but noted that the approach would only work if over- or under-reporting, or other issues with landings and removals data were similar over the entire time series.

In addition to the issues of estimating landings and removals for more recent years, the AT also needed a method to estimate landings during the historical period, defined as the period from a year in which the population could be considered to be at virgin levels, to the time at which landings data become available. The AT addressed this issue by assuming that fishing effort increased during the historical period, and explored the effects of the assumed nature of this increase (e.g. linear versus exponential) on the assessment results using additional model runs with different assumptions. The RP accepted that this was a reasonable approach, but also agreed with the AT that there was considerable uncertainty about the removals during the historical period and therefore the status of the populations at the time when landings data became available.

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

Two assessment methods were used across the four stock assessments, both being variants on the basic Age-Structured Production Model. For sandbar and the two blacknose stocks a state-space variant was used: the State-Space Age-Structured Production Model (referred to here as ASPM); while for the dusky assessment the catch-free variant was used: Age-Structured Catch-Free Production Model (ASCFM).

Overall, it is the conclusion of the RP that these approaches are appropriate to the assessment of these stocks given some of the particular issues that relate to shark assessments, e.g., uncertain catch histories, and low productivity. Nevertheless there is improvement that could be made to the assessment platforms. Specific comments on each method and the applications in the individual assessments are provided in the sections below.

The RP notes that the AT had hoped to use Stock Synthesis (SS) instead of the variant on the age-structured production model for the three stocks assessed with ASPM. It is the understanding of the RP that at this time SS does not include the specific parameterization of some of the key biological processes used in these shark assessments, which are available within ASPM. Further we note that given the resources and time available to the AT—in particular that four assessments were to be completed—it would have been difficult for them to develop SS models for these stocks. While SS would have allowed the separation of the sexes, and estimation of length-specific selectivity, it is not presently known if setting up the model in this way, as would have been preferred by the RP, would have markedly changed the assessment results.

ASPM – Sandbar and Blacknose Assessments

The state-space age-structured production variant of the ASPM model was applied to the assessments of sandbar sharks and the two blacknose shark stocks. It has previously been applied to shark stocks, including some of the stocks being assessed here.

The basic model is quite simple – a combined or unisex age-structured model with recruitment modeled via a Beverton-Holt spawner recruitment curve and selectivity and other biological processes (e.g. natural mortality and maturity) modeled as age-specific processes. In particular ASPM incorporates pup survival and density-dependence, important biological characteristics of shark populations, which makes it easy to explicitly incorporate information on the productivity of the stocks which is a key constraining factor in shark stock assessment. ASPM requires the assumption that at the start of the model period the population was in a virgin state, and in the implementations for the three stocks recruitment was estimated as deterministic and selectivity at age was fixed.

Whilst ASPM was considered appropriate for the assessments undertaken, the RP had some questions as to the additional benefit that came from the additional complexity allowed through the state-space implementation. Briefly, in the three assessments the state-space component was implemented as a random walk on deviates on the fishery-specific effort trajectories. Effort was then used to predict catch through other model quantities, e.g. abundance and catchability. So catches are not assumed to be known without error and therefore the model trades off the penalties on the effort deviates with the CVs on the catch estimates. This is essentially a random walk in fishery-specific fishing mortality. This approach requires the estimation of a large number of additional parameters, although they are constrained deviates so that the effective number of estimated parameters is less than the absolute number.

A second issue of the ASPM relates to the weighting of different data sources – a very well-known important issue in stock assessments. ASPM has incredible flexibility in the weighting of catch and CPUE data – through time specific CV's on individual catch and CPUE series ($w_{i,j}$); lambda scaling factors on individual series, and overall model CVs. When you add to this the process error variances on the effort deviates it made it difficult for both the AT and the RP to determine the actual weightings that were being applied/estimated for the various data sources. This was particularly important in determining the weightings provided to the different abundance indices when trying to understand the fit to some of these series. The use of lambdas may also have theoretical implications for estimates of parameter uncertainty.

In the sandbar assessment, where an assumption was made to fit the catch data five times better than the CPUE indices, it essentially resulted in the assumption that catches were known without error and all the complexity with the state-space implementation and various lambdas just made things more difficult to understand. However, in the case of the GoM blacknose assessment where the catch and CPUE were given equal weight – the lack of fit to the shrimp trawl catches actually provided some insights into the tensions that were going on in the model between the catch series, key CPUE series, and productivity.

With the sandbar assessment, the RP identified problems in fits to early catch data due to a mismatch between catch series and the different assumptions of the 'historic' and 'modern' periods. This was another example of where the added complexity of the modeling made the implementation of the assessment harder as there were more model options that are needed to be set.

With the GoM blacknose assessment problems in the implementation of the state-space component meant that they were unable to get a satisfactory model with different selectivities for the shrimp bycatch fishery pre- and post- TED implementation. So the post-TED selectivity was assumed for the entire model period. The RP requested a model run with a selectivity curve that might better approximate the pre-TED situation and the results were broadly similar, so the problem encountered in the assessment probably did

not have a qualitative impact on the results. Nevertheless it is another example where the additional complexity did not necessarily help in the assessment process.

Principal data inputs for the blacknose GoM assessment were historical catches and the abundance indices. Evidence of the acceptability of the assessment depends in particular on how well the model was able to fit to the input data. The abundance indices generally either showed no trend or an increasing trend over recent years – particularly for those indices given a high ranking by the DW (NFMS SE LL, SEAMAP summer and fall and SEFC shark BLL OB indices). Sensitivity runs requested by the RP demonstrated that the blacknose GoM assessment model was unable to fit apparent trends in the abundance indices at all, unless implausible additional historical catches were also estimated. This fundamental lack of fit of the model to the input data caused the RP to reject the blacknose GoM assessment model. A remedy for the situation would involve the development and application of a model with additional but plausible flexibility (e.g. in perhaps annual recruitment variation) to provide improved fits to observation data.

Some other secondary comments on the use of ASPM for the three assessments which should be considered in either the interpretation of the results and/or the consideration of modeling approaches for future assessments:

- Assuming fixed values or tight priors on pup survival, with other biological parameters fixed (e.g. fecundity and natural mortality) implies very tight, or exact estimates for key population productivity parameters such as the annual number of replacement spawners per spawner at low population size (α), which translates into steepness and F_{MSY}. This is not necessarily a bad thing – we know that the productivity of sharks is low, but people should be aware that the assumptions that go into the model tightly constrain some key reference points. The population trends were typically a 'one-way trip' so the data are unlikely to be that informative with respect to population productivity. This is discussed further under ToR 4.

- Both ASPM and the ASCFM removes catches starting from age 1 (depending on the selectivity), but there were several instances where size data suggested major catches of age-zero individuals. The problem with extending the catches back to age zeros relates to the assumption that it is at this stage where density dependence occurs, e.g. pup survival increases at lower stock sizes. Both the assumption that pup survival is the source of density dependence and the assumption that age zero catches are zero should be addressed in future assessments. In the current implementation, the use of catch in numbers meant that instead removing these fish at age 1 probably did not have a significant impact on the assessment results – but this might not hold if the fishing of age zeros was prior to density dependence.

- For all assessments the selectivity curves were estimated/guesstimated external to the modeling process – this is clearly not ideal and discussed under ToR 1. Under this ToR the RP also discussed the problem of the age-length key and the use of a single age-length key to convert length to age over the entire period for which length samples are available. All of these factors suggest that converting length data to age data and estimating age-based selectivity is not ideal. The RP recommends that future modeling approaches fit to catch-at-length data and, if possible, estimate length-specific rather than age-specific selectivity curves.

- The assumption of initial virgin conditions was unlikely to be met, but the absence of early age composition data makes it difficult to try and estimate the initial levels of depletion. The RP requested various runs for the ASPM assessments in order to investigate this assumption.

- Based on discussions during the review it was determined that the AT experienced problems obtaining convergence (as measured by obtaining a positive definite Hessian matrix) with several of the models for some of the species. With these models that did not converge they also found instances where alternative starting values led to different results and that this commonly occurred with the effort deviates. In order to assist in future assessment reviews tables of parameter values and likelihood components by

estimation phase should be provided and specific sensitivity analyses to alternative starting values should be undertaken for key model runs. This is covered in more detail under ToR 7.

In order to address the various issues described in this section, numerous sensitivity analyses were undertaken by the AT, including several at the request of the RP.

So on balance, the RP concludes that the general population dynamics assumed in the assessment model are appropriate for the assessments, but it is not clear to the RP that the added complexity of the state space implementation of ASPM was useful in these assessments. This complexity includes the use of so-called 'historic' and 'modern' periods where different assumptions are made about the reliability of catch data and the relationship between catch and effort. Importantly, the RP believes that using a model that did not include this added complexity would not change the conclusions of the assessments or this review, but would only have made the models and model results easier to interpret, in particular allowed for easier analysis of the impacts of alternative data weighting scenarios.

ASCFM – Dusky Shark

The underlying population dynamics for ASCFM are essentially the same as those for ASPM – the major difference in the approaches being that absolute estimates fishery removals are not used within ASCFM. This is the same modeling platform used for the previous dusky assessment and is necessary due to the difficulties in obtaining any sort of catch history.

In the current implementation of ASCFM, effort series for the various fleets were used as input to the model and fishing mortality modeled as a function of effort. A random walk in fishing mortality is estimated through time, and specific to this assessment, a break point in F at the year 2000 was incorporated into the model to allow for significant change in management that occurred at that time (the ban on retention of dusky sharks).

In the case of a single fleet the relationship between effort and overall fishing mortality is relatively simple, but in this implementation multiple fleets were exploiting the stock. Estimates of relative effort for the fleets were determined. The actual quantity that is of interest would be the 'effective effort' of each fleet and this requires the incorporation of the catchability of each fleet, e.g. how many dusky sharks are caught per unit of effort for each fleet. One improvement to future dusky assessments will be the incorporation of information on the relative catchability of the different fleets and this could be done by comparing the ratios of CPUE for the different methods, with the same units of effort. Preferably this is done over a period when the fleets overlap and it is important to consider any differences in selectivity of the fleets when making such comparisons.

One concern of the RP was the lack of simulation studies to determine the performance of the ASCFM under known conditions, particularly the response of the model to alternative scenarios about productivity when catches are known. This is further discussed under ToR 9.

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

The RP accepted the model results for Atlantic blacknose, sandbar and dusky sharks, but not GoM blacknose sharks. There was considerable uncertainty in the accepted results that the RP considers should be conveyed as part of the RW recommendations. The AT had assigned a base case model in each assessment as a suitable basis for conducting sensitivity analyses, but not with the interpretation that it provided a central best case for the provision of management advice. Dimensions explored by the AT in sensitivity analyses were not necessarily as upper and lower dimensions from the base case, and often only as a plausible alternative to the base, with unknown relative probability. Given the general uncertainty in these shark assessment inputs and results, the RP was unable to recommend central best estimates for stock abundance, biomass and exploitation values.

The RP considered that major dimensions in shark assessment uncertainty were in the overall biomass scale and the stock productivity. A data input that affects biomass scale is the estimated level of historical catch, and productivity is particularly affected by pup survival, reproduction periodicity, and natural mortality. Plausible sensitivity scenarios were developed by the RP in consultation with the AT to bracket the uncertainty in biomass scale, productivity, and other important dimensions particular to each shark species (Table 2). Sensitivities can be considered as plausible alternative states of nature of the stock for the interpretation of population benchmarks, stock status and management recommendations. For Atlantic blacknose the RW-base case differed from the base case in the assessment report in that the UNC series was down-weighted to better account for higher uncertainty in that index relative to others (rank 4 by the data workshop), and poor model fit to the UNC index.

Blacknose Shark: GoM Stock

For the GoM blacknose shark stock, there were two main issues that were not sufficiently reconciled at the RW, and as a result, the RP in unwilling to accept that stock status can be determined based on the RW analyses. The first of these is the uncertainty in the status of the population at the start of the historical period, when the population is assumed to be at a virgin size. In the case of GoM blacknose shark, the shrimp bycatch comprises most of the catches, and this fishery existed before the start of the historical period (1950). Although it is not known whether bycatch levels would have been similar in the past, this assumption is difficult to justify, but statements about the status of relative to a biomass benchmark are based on this assumption. The second issue pertains to difficulties fitting to both the catches and the survey indices simultaneously. The catch series shows relatively stable catches until about 2005 followed by gradual decline. Given the low productivity of the stock, when these catches are reasonably fit, the model estimates a general downward trend in abundance from 1950 to about 2008. In contrast, the BLLOP, NMFS SE LL, SEAMAP summer and SEAMAP fall indices appear to indicate stable or increasing abundance trends and the marked residual patterns indicate how poorly the model results fits these indices. At the RW, the AT did a model run with a very low weight on the catch data in order to see what the predicted catch series would look like if the indices were fit well. Both the magnitude and trend of the predicted catches were sufficiently different from the observed catches, that it was not possible to reconcile the catch and abundance index time series at the RW.

Blacknose Shark: Atlantic Stock

The unit used for spawning biomass was spawning stock fecundity (SSF: the female numbers multiplied by the age-specific fecundity). Current spawning biomass depletion was estimated to be between 0.17 and 0.26 across the range of sensitivities (SSF_{2009}/SSF_0, Table 3). Current fishing mortality on the stock was estimated to be between 0.29 and 0.48 per year in 2009 (F_{2009}, Table 3). Trajectories of SSF/SSF_{MSY} and F/F_{MSY} through time show a fairly consistent pattern across all sensitivities (Figure 1). Spawning biomass generally declined from about 1986 to 2004, and flattened off in recent years. Fishing mortality generally increased in the period from about 1990 to 2000, then reduced to a low level by 2006, and again increased to levels comparable to previous maximums by 2009. Abundance estimates in 2009 range from about 107,000 to about 439,000 sharks (Table 3), and fishing mortality estimates in 2009 range from 0.29 to 0.48.

Sandbar Shark

Current spawning biomass depletion was estimated to be between 0.18 and 0.34 across the range of sensitivities (SSF_{2009}/SSF_0, Table 4). Current fishing mortality on the stock was estimated to be between 0.01 and 0.02 per year in 2009 (F2009, Table 4). Trajectories of SSF/SSF_{MSY} and F/F_{MSY} through time show a fairly consistent pattern across all sensitivities (Figure 3). An exception was the RW-4 low productivity scenario that showed considerably increased F/F_{MSY} values in comparison with the other sensitivities. Spawning biomass generally declined from about 1986 to 2007, and then showed a slight

increase to 2009. Fishing mortality varied about a constant level of about 0.08 (apical F) from 1983 to 2003, and then declined to below 0.03 by 2009. Abundance estimates in 2009 range from just less than 1 million to about 4.6 million sharks.

Dusky Shark

Current spawning biomass depletion was estimated to be between 0.13 and 0.24 across the range of sensitivities (SSB2009/SSB0, Table 5). Current fishing mortality on the stock was estimated to be between 0.026 and 0.080 per year in 2009 (F2009, Table 5). Trajectories of SSB/SSBMSY and F/FMSY through time show a fairly consistent pattern across all sensitivities (Figure 5). An exception was the S18 low productivity scenario that showed considerably increased F/FMSY values in comparison with the other sensitivities. Spawning biomass generally declined from about 1980 to 2009. Fishing mortality increased during the period 1982 to 1999, declined to a comparably low level by 2008, and then either flattened off or increased slightly in 2009.

Table 2. Sensitivity analyses selected for the RW.

Run	Code	Description
Blacknose Atlantic		
Base	RW-Base	Base case as provided by the AT with down-weighted UNC index
Inverse CV	RW-S1	Inverse CV abundance index weighting
1 year cycle	RW-S2	One year reproduction cycle
High catch	RW-S3	Catch increased one standard deviation
Low catch	RW-S4	Catch decreased one standard deviation
High productivity	RW-S5	Fecundity fixed at 6 pups for all ages, pup survival increased to 0.90
Low productivity	RW-S6	Pups per female reduced to 1, pup survival reduced to 0.75, M for ages 1-max increased to 0.25
Sandbar		
Base	Base	Base case as provided by the AT
Inverse CV	S1	Inverse CV abundance index weighting by the AT
2 year cycle	S5	Two year reproduction cycle by the AT
3 year cycle	S6	Three year reproduction cycle by the AT
High catch	RW-S1	Midpoint of base and high catch scenario of S13 by the AT
Low catch	RW-S2	Midpoint of base and low catch scenario of S12 by the AT
High productivity	RW-S3	Fecundity fixed at 9.5 pups for all ages, pup survival increased to 0.90, *M* for ages 1-max set to 0.105
Low productivity	RW-S4	Pup survival reduced to 0.80, *M* for ages 1-max increased by 10%
Dusky		
Base	Base	Base case as provided by the AT
High *M*	S3	Base *M* multiplied by 1.342
U-shaped *M*	S4	Elevated *M* for older age classes
High productivity	S17	Pups per female 10, two year reproductive cycle, pup survival 0.97
Low productivity	S18	Pups per female 4, pup survival 0.51

Table 3. Results of scenarios selected to explore the range of model outputs for Atlantic blacknose shark.

	RW-Base		**RW-S1 (Inv-CV)**		**RW-S2 (1-yr cycle)**		**RW-S3 (high catch)**		**RW-S4 (low catch)**		**RW-S5 (high productivity)**		**RW-S6 (low productivity)**	
	Estimate	**CV**	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**
SSF_{2009}/SSF_{MSY}	**0.60**	0.16	**0.43**	0.16	**0.61**	0.18	**0.58**	0.16	**0.64**	0.16	**0.61**	0.18	**0.55**	0.15
F_{2009}/F_{MSY}	**5.02**	0.32	**4.77**	0.36	**3.37**	0.32	**5.51**	0.33	**4.67**	0.32	**3.26**	0.32	**22.53**	0.32
SSF_{MSY}/SSF_0	0.41		0.41		0.33		0.41		0.41		0.33		0.47	
MSY	24495		22978		20810		66625		17910		20429		36996	
SPR_{MSY}	0.67	0.03	0.67	0.04	0.48	0.04	0.67	0.03	0.67	0.03	0.46	0.04	0.94	0.034
F_{MSY}	0.08		0.07		0.14		0.08		0.08		0.15		0.01	
SSF_{MSY}	96809		90814		123900		288360		77577		116650		104620	
N_{MSY}	153709		144550		122172		576722		155385		118788		247916	
F_{2009}	0.38	0.32	0.34	0.36	0.46	0.32	0.41	0.33	0.35	0.32	0.48	0.32	0.29	0.32
SSF_{2009}	58049	0.19	38816	0.17	76066	0.20	168300	0.19	49395	0.19	71346	0.20	57920	0.19
N_{2009}	155000		107418		120381		439136		131490		116155		222969	
SSF_{2009}/SSF_0	0.24	0.08	0.17	0.11	0.21	0.19	0.24	0.08	0.26	0.07	0.20	0.17	0.26	0.14
B_{2009}/B_0	0.22	0.17	0.16	0.14	0.20	0.18	0.21	0.15	0.24	0.15	0.20	0.19	0.22	0.16
R0	85148	0.06	79571	0.08	66366	0.06	252780	0.07	68012	0.06	64308	0.06	145330	0.06
Pup-survival	0.81		0.81		0.81		0.81		0.81		0.90		0.75	
alpha	2.26		2.26		4.52		2.26		2.26		5.02		1.14	
steepness	0.36		0.36		0.53		0.36		0.36		0.56		0.22	

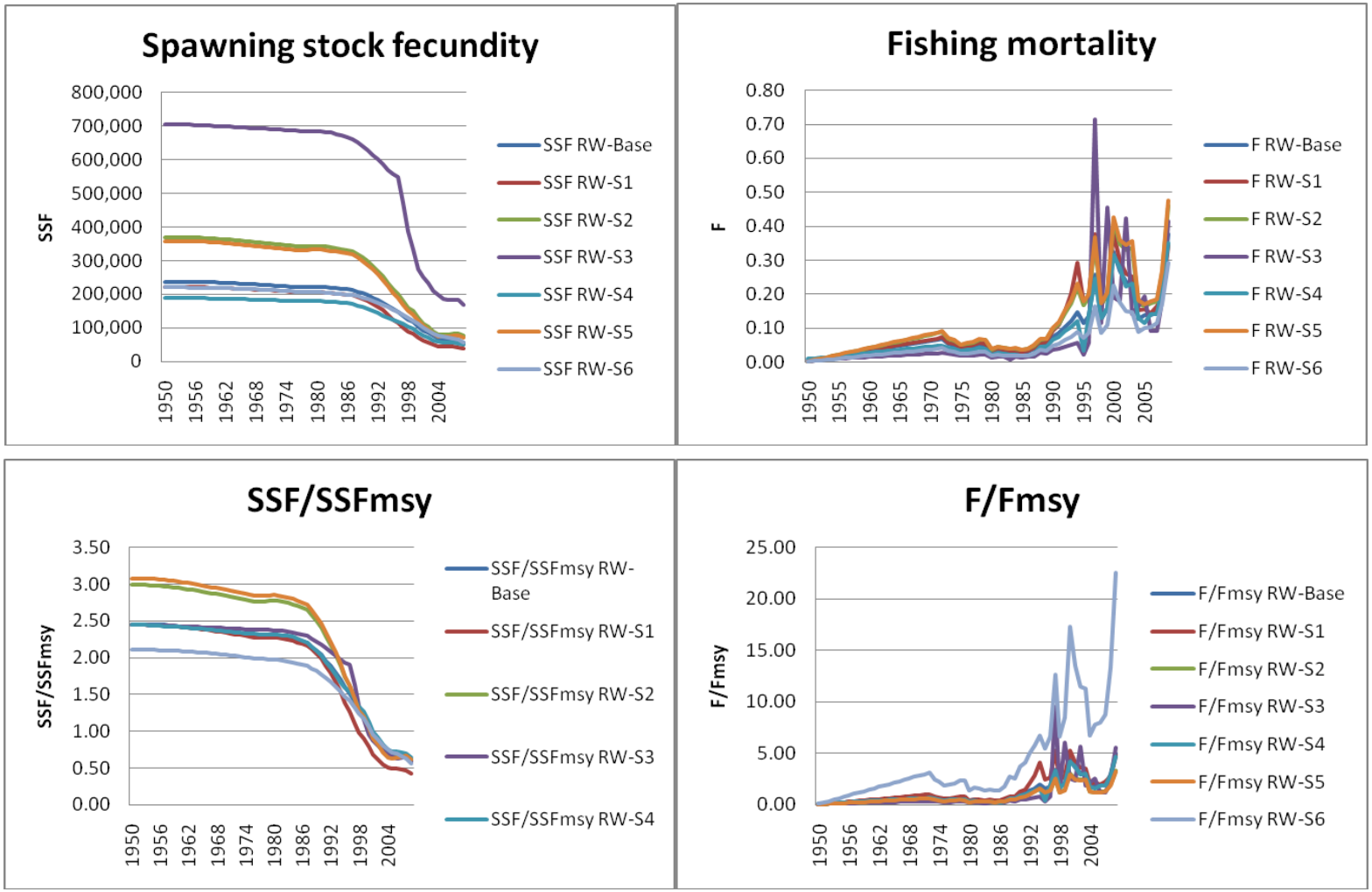

Figure 1. Time trajectories of key stock status indicators for Atlantic blacknose shark. Four trajectories are shown: SSF (spawning stock fecundity; top left panel), total apical F (top right panel), relative biomass (bottom left panel), and relative fishing mortality (bottom right panel). Each line within a panel is a different sensitivity analysis.

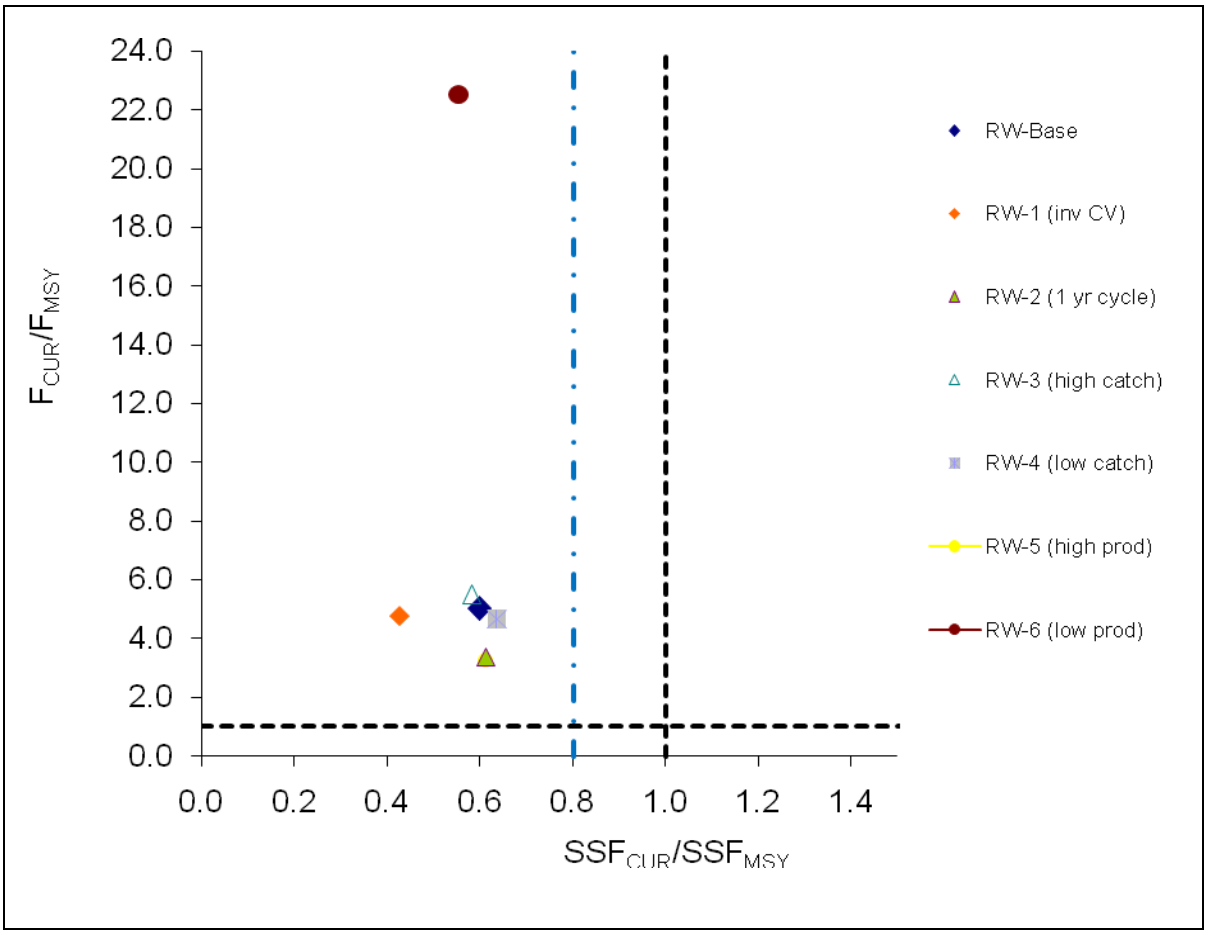

Figure 2. Phase plot of stock status for Atlantic blacknose shark, MSST shown as blue vertical dashed line.

Table 4. Results of scenarios selected to explore the range of model outputs sandbar shark.

	BASE		S5 (2 yr rep cycle)		S6 (3 yr rep cycle)		S1 (Inv CV)		RW-1 (high catch)		RW-2 (low catch)		RW-3 (high prod)		RW-4 (low prod)	
Parameter	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**	**Estimate**	**CV**
AICc	718.01		717.81		718.71		652.84		715.02		716.89		716.91		716.16	
Objective function	117.95		117.85		118.30		85.37		116.46		117.39		117.40		117.02	
SSF_{2009}/SSF_{MSY}	**0.66**	0.83	**0.64**	0.71	**0.66**	1.09	**0.56**	0.70	**0.71**	0.78	**0.67**	0.85	**0.51**	0.55	**0.72**	3.49
F_{2009}/F_{MSY}	**0.62**	0.57	**0.46**	0.55	**0.93**	0.61	**0.62**	0.44	**0.29**	0.57	**0.64**	0.57	**0.36**	0.48	**2.62**	0.61
N_{2009}/N_{MSY}	0.74		0.74		0.78		0.65		0.77		0.76		0.87		0.70	
MSY	160643		152940		173414		152907		461238		118699		98928		194389	
SPR_{MSY}	0.78	0.06	0.69	0.09	0.86	0.04	0.74	0.09	0.78	0.06	0.77	0.07	0.53	0.14	0.95	0.01
F_{MSY}	0.021		0.030		0.030		0.025		0.022		0.019		0.059		0.004	
SSF_{MSY}	477590		503420		503420		430320		1377800		349330		425530		530410	
N_{MSY}	1928165		1768504		2012907		1804687		5530573		1427463		1037329		2500141	
F_{2009}	0.01	0.57	0.01	0.55	0.01	0.61	0.02	0.44	0.01	0.57	0.01	0.57	0.02	0.48	0.01	0.61
SSF_{2009}	312890	0.60	319760	0.59	313510	0.63	240950	0.40	984770	0.58	234320	0.60	215900	0.55	381620	0.61
N_{2009}	1539102		1408804		1688767		1277408		4605900		1165723		975580		1899533	
SSF_{2009}/SSF_0	0.28	0.41	0.25	0.42	0.32	0.41	0.24	0.27	0.32	0.38	0.28	0.41	0.18	0.44	0.34	0.38
B_{2009}/B_0	0.34	0.33	0.33	0.33	0.35	0.34	0.30	0.18	0.37	0.31	0.34	0.33	0.27	0.33	0.36	0.33
R0	563490	0.20	516810	0.18	612140	0.23	516900	0.14	1587000	0.21	423250	0.20	281740	0.12	774030	0.24
Pup-survival	0.84	0.29	0.84	0.29	0.84	0.29	0.94	0.30	0.84	0.29	0.85	0.29	0.90	0.29	0.76	0.29
alpha	1.64		2.05		1.37		1.84		1.65		1.66		3.80		1.10	
steepness	0.29		0.34		0.25		0.31		0.29		0.29		0.49		0.22	
SSF_0	1097900	0.20	1258700	0.18	993980	0.23	1007200	0.14	3092300	0.21	824700	0.20	1192200	0.12	1121000	0.24
SSF_{MSY}/SSF_0	0.43		0.40		0.48		0.43		0.45		0.42		0.36		0.47	

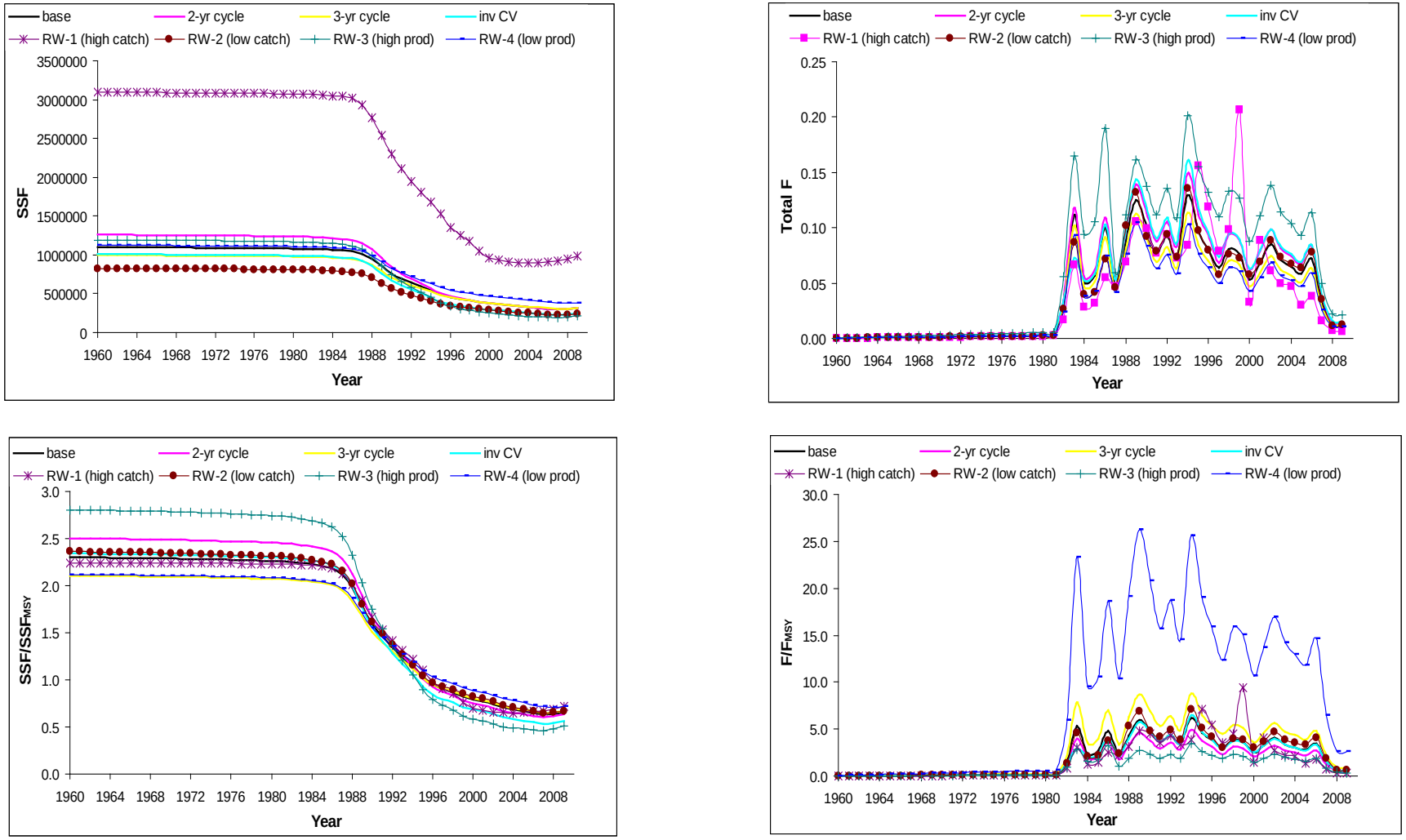

Figure 3. Time trajectories of key stock status indicators for sandbar shark. Four trajectories are shown: SSF (spawning stock fecundity; top left panel), total apical F (top right panel), relative biomass (bottom left panel), and relative fishing mortality (bottom right panel). Each line within a panel is a different sensitivity analysis.

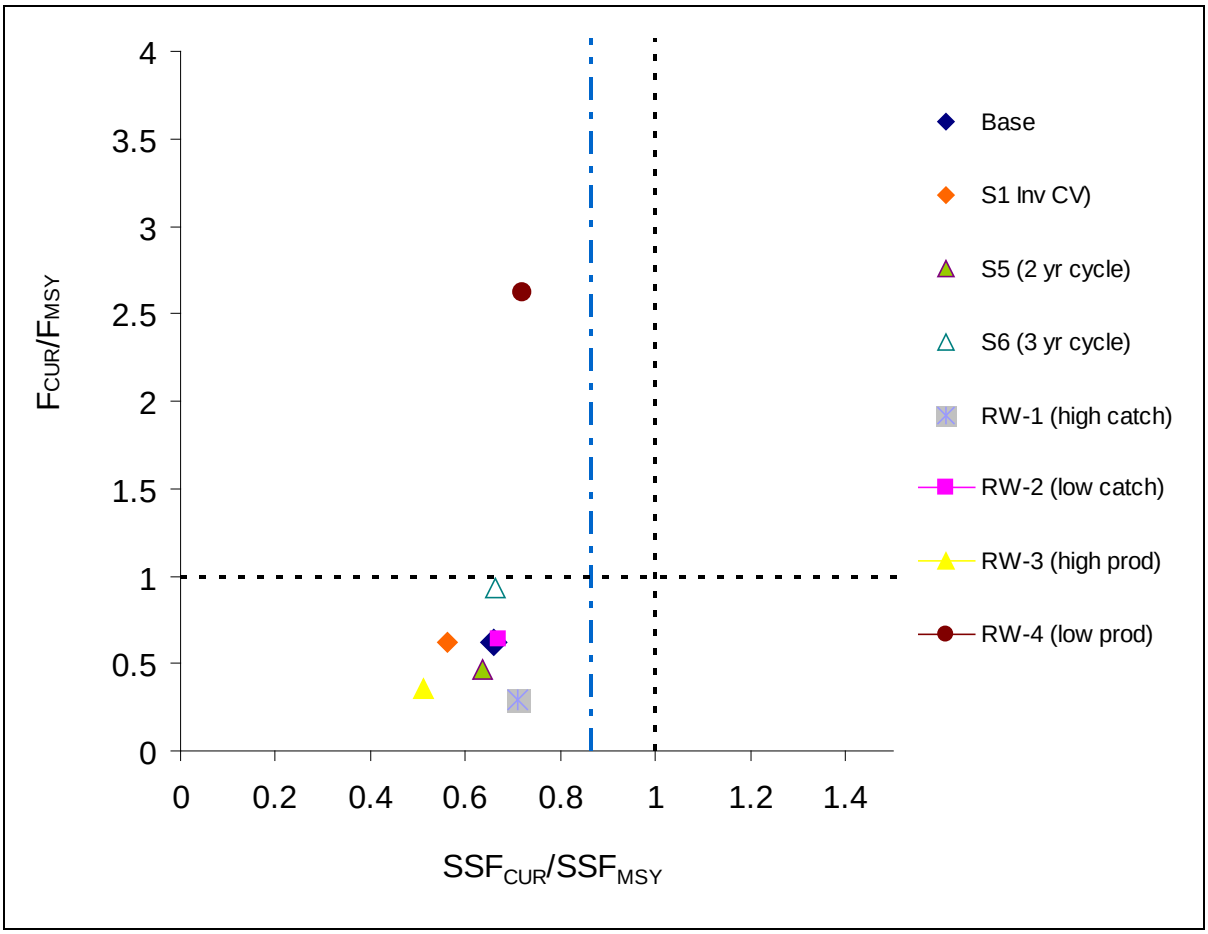

Figure 4. Phase plot of stock status for sandbar shark, MSST shown as blue vertical dashed line.

Table 5. Results of scenarios selected to explore the range of model outputs for dusky shark.

Run	**Base**	**S3**	**S4**	**S17**	**S18**
Description	--	High M	U shaped M	High productivity	Low productivity
F_{2009}	0.054	0.034	0.026	0.080	0.030
F_{MSY}	0.035	0.017	0.019	0.054	0.007
SSB_{2009}/SSB_0	0.15	0.18	0.18	0.13	0.24
SSB_{MSY}/SSB_0	0.35	0.43	0.43	0.28	0.47
$\mathbf{SSB_{2009}/SSB_{MSST}}$	**0.46**	**0.45**	**0.44**	**0.49**	**0.53**
$\mathbf{SSB_{2009}/SSB_{MSY}}$	**0.44**	**0.42**	**0.41**	**0.45**	**0.5**
$\mathbf{F_{2009}/F_{MSY}}$	**1.55**	**2.01**	**1.39**	**1.49**	**4.35**
Pup survival	0.89	0.95	0.96	0.97	0.51
Steepness	0.51	0.32	0.32	0.71	0.25

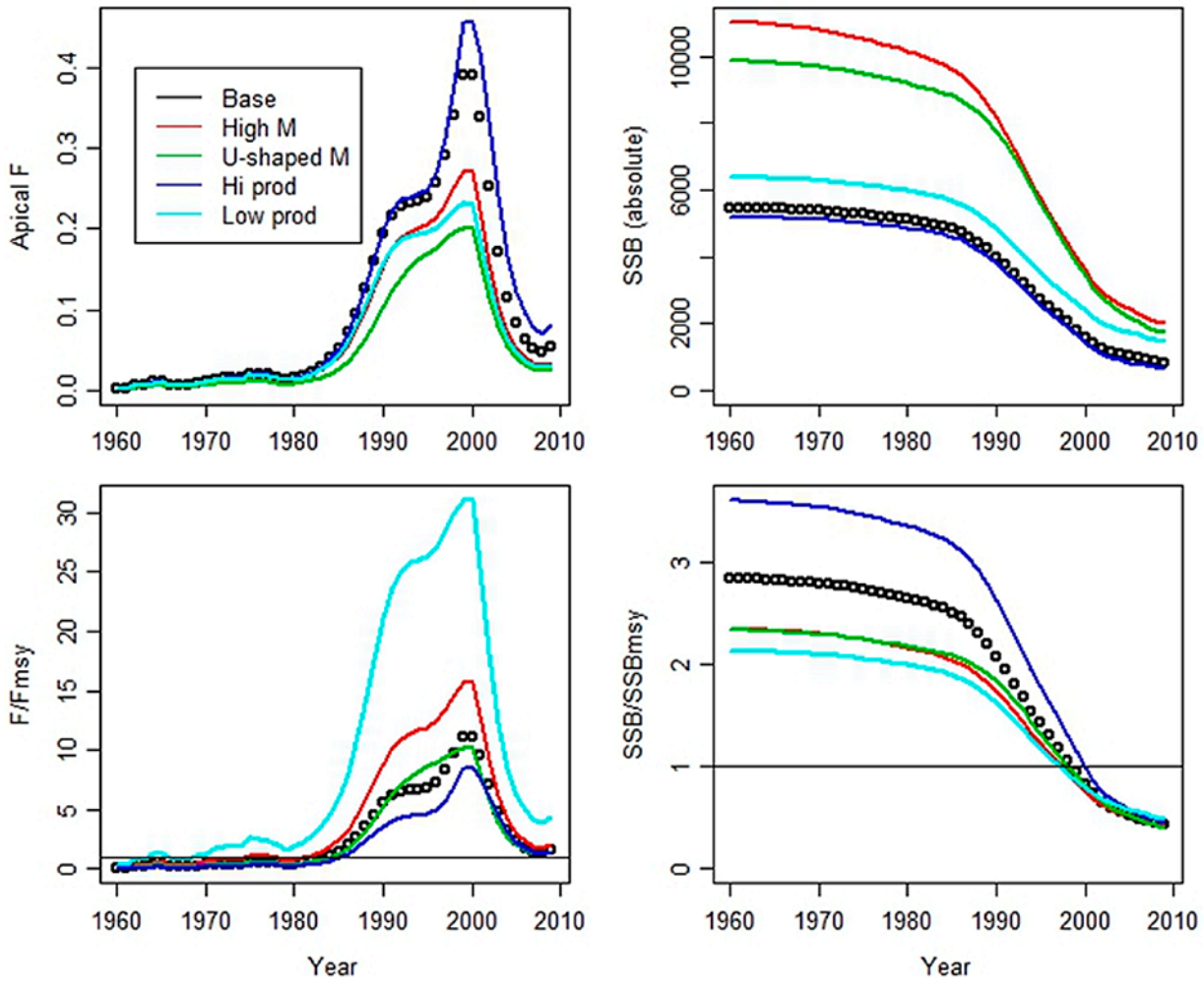

Figure 5. Time trajectories of key stock status indicators for dusky shark. Four trajectories are shown: total apical F (top left panel), SSB (spawning stock biomass; top right panel), relative fishing mortality (bottom left panel) and relative biomass (bottom right panel). Each line within a panel is a different sensitivity analysis.

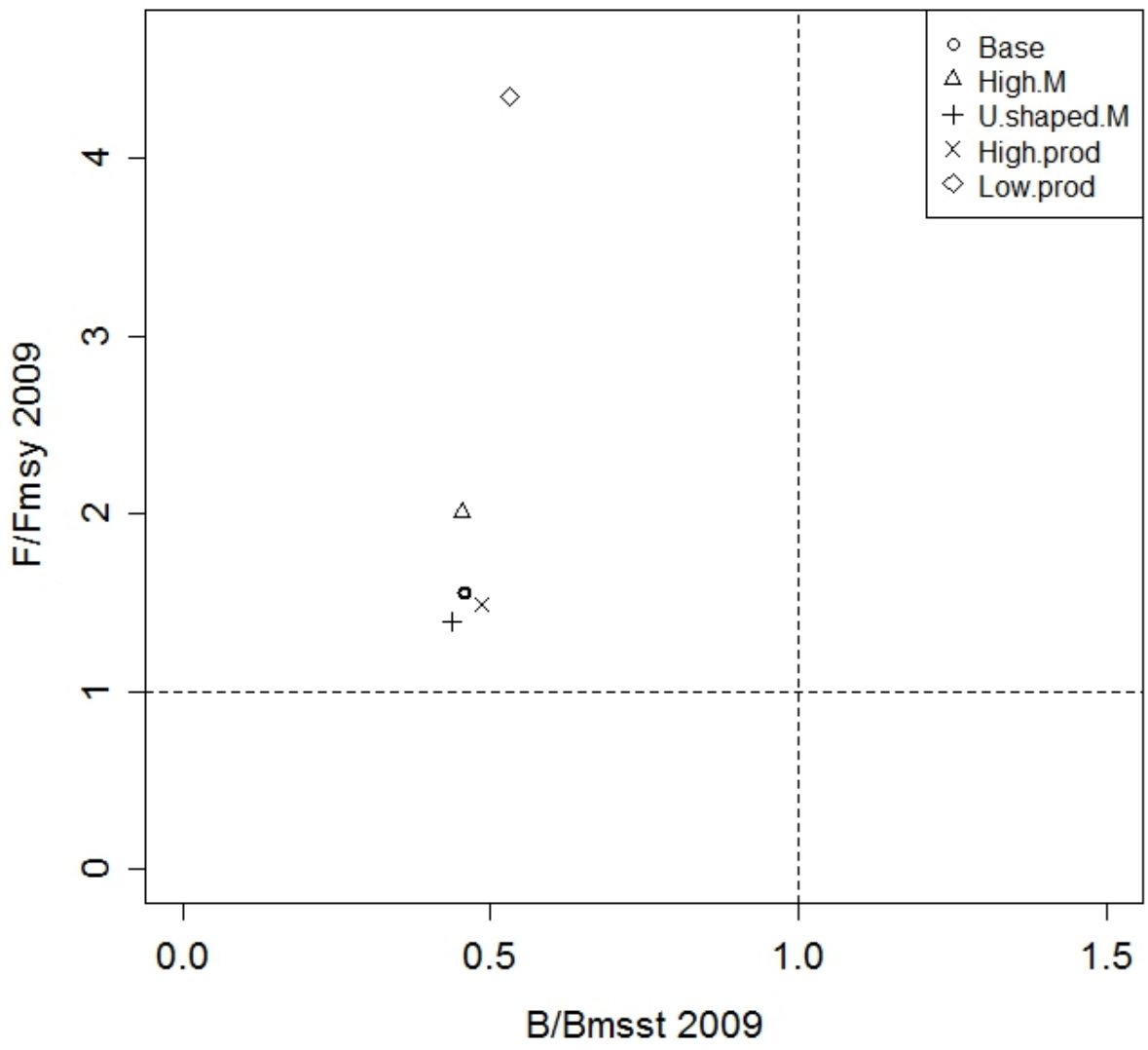

Figure 6. Phase plot of stock status for dusky shark (note that BMSST is used as a denominator for the X axis and not B_{MSY} as for the other species).

ToR 4: Evaluate the methods used to estimate population benchmarks and stock status (e.g., MSY, FMSY, BMSY, MSST, MFMT, or their proxies); recommend appropriate management benchmarks, provide estimated values for management benchmarks, and declare stock status, consistent with the stock status determination criteria, benchmark, and biological reference points in the Consolidated HMS FMP, proposed FMPs and Amendments, other ongoing or proposed management programs, and National Standards.

Methods used to calculate population and management benchmarks were appropriate and followed guidelines provided by Restrepo et al. 1998 and procedures developed for the 2006 assessments (Brooks et al. 2010). Values of MSY, FMSY and SSFMSY were estimated by the assessment models (special case for dusky shark). The minimum spawning stock size threshold (MSST) was defined as [(1-M) or 0.5 whichever is greater]*SSFMSY, the default for data-moderate situations in the guidelines. The maximum fishing mortality threshold (MFMT) was defined as FMSY.

The method for the calculation of MSST was made more complicated because of the age-specific M used for these shark species. An age-independent average M value that results in the same survivorship of fish to the plus group was used for MSST calculations.

Table 6. Principal stock indicator ranges over the RW sensitivities per shark species.

	Blacknose Atlantic	**Sandbar**	**Dusky**
SSF_{2009}/SSF_{MSY}*	0.43 - 0.64	0.51 - 0.72	0.41 - 0.50
F_{2009}/F_{MSY}	3.26 - 22.53	0.29 - 2.62	1.39 - 4.35

* Dusky shark used spawning stock biomass (SSB) rather than spawning stock fecundity (SSF).

Atlantic Blacknose Shark

Results over the alternative sensitivity analyses all showed that the stock was overfished (SSF_{2009}/SSF_{MSY} of 0.43 to 0.64, all below MSST) (Figure 2, Table 6) and therefore subject to rebuilding. Current F values over all sensitivities also indicated that the stock was subject to overfishing (F_{2009}/F_{MSY} 3.26 to 22.53).

Sandbar Shark

Results over the alternative sensitivity analyses all showed that the stock was overfished (SSF_{2009}/SSF_{MSY} of 0.51 to 0.72, all below MSST) (Figure 4, Table 6) and therefore subject to rebuilding. Current *F* values over most sensitivities indicated that the stock was not currently subject to overfishing (F_{2009}/F_{MSY} 0.29 to 0.93). However, the low productivity scenario did indicate overfishing (F_{2009}/F_{MSY} 2.62)

Dusky Shark

Results over the alternative sensitivity analyses all showed that the stock was overfished (SSB_{2009}/SSB_{MSY} of 0.41 to 0.50) (Table 6, Figure 6) and therefore subject to rebuilding. Current *F* values over all sensitivities also indicated that the stock was subject to overfishing (F_{2009}/F_{MSY} 1.39 to 4.35).

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

For the four stocks being assessed projection analyses are very important, in particular the propagation of uncertainty in the projections. This is for two reasons 1) these stocks are currently all subject to rebuilding plans; and 2) some of the projection benchmarks rely not on the median of the projections, rather some percentile of the distribution.

Two approaches were used for undertaking projections and the split was the same as that for the assessment methods. For the stocks assessed using ASPM (sandbar, and the two blacknose stocks) projections were carried out using Pro-2Box. The sandbar shark assessment report describes the procedure as it was applied in that assessment and we have repeated the important elements below:

"Projections were bootstrapped ≥500 times by allowing for process error in the spawner-recruit relationship. Lognormal recruitment deviations with SD = 0.4, with no autocorrelation, were assumed. No other variability was introduced into the projections. Under these assumptions, the base model was projected at F = 0 to determine the year when the stock can be declared recovered with a 70% probability (SSF/SSFMSY > 1). If that year is >10, then management action should be implemented to rebuild the stock within the estimated rebuilding time+1 generation time."

An alternative approach was used for projections for the dusky shark assessment and is detailed below:

"Projections were governed with the same set of population dynamics equations as the original assessment model, but allowed for uncertainty in initial conditions at the beginning of the time series (that is, in 2009) as well as in underlying productivity. Projections were run using Monte Carlo bootstrap simulation, where initial biomass (B_{2009}^{boot}), fishing mortality (F_{2009}^{boot}), and pup survival at low biomass ($\exp(-M0)_{2009}^{boot}$) were sampled from a multivariate normal distribution with expectations equivalent to posterior modes from the base run, and standard deviations set to the posterior standard deviation (obtained numerically by rejection sampling of the "profile likelihood" posterior approximation). Covariance values were obtained from the Hessian approximation of the variance-covariance matrix at the posterior mode. The multivariate normal approximation was chosen because it reduces the probability of selecting values of the different parameters that are unlikely to have generated the data (for instance, high fishing mortality and low pup survival)."

The key difference between the two approaches is that the latter method considers uncertainty in two key additional model quantities, the abundance of the stock in the terminal year and the level of productivity. Further it incorporates the correlation that exists between these estimated quantities.

Given the critical importance of incorporating uncertainty in the projections, the RP did not accept the projections or projection method applied to the sandbar and south Atlantic blacknose stocks. However, although not all sources of variability are included in the projections, it is the opinion of the RP that the method applied to the dusky assessment is sufficient given the greater uncertainties in different scenarios.

The RP recommends that projections be undertaken for sandbar and south Atlantic blacknose stocks using a method similar to that applied to dusky shark. The projection methodology should at least:

- Incorporate uncertainty in the overall abundance estimate in the terminal year;
- Incorporate uncertainty in the key productivity parameters, if estimated;
- Incorporate any correlation in the estimation of the above quantities; and
- Incorporate low levels of stochasticity in future recruitment – consistent with the tightly constraining biology of the species.

In making this recommendation, the RP was not aware of a generic package available to do the projections and that it was not possible to complete these projections within the RW.

Even with a more accurate reflection of the within model uncertainty being propagated through the projections, the RP was of the opinion that it was necessary to carry through the structural uncertainty considered in bracketing the current stock conditions (ToRs 3 and 4) through into the projections.

The RP prepared the following set of tables to contain the projection results. The key aspects of this table are:

- That multiple scenarios or possible ‘states of nature’ are included;
- Current terminal F’s are included to allow comparisons across runs and to allow examination in the relative change in F necessary to achieve the particular rebuild strategies;
- It includes projection scenarios requested by HMS that relate to the current management arrangements;
- It includes most of the results for dusky shark requested by the RP – the others will be done after the meeting; and
- It has no results for sandbar and south Atlantic blacknose because as noted above a satisfactory package to undertake the projections was not available to the AT during the RW.

When considering results similar to those presented in the tables below and developing a management response it is often useful to look not only at the results for particular scenarios, but also to examine the results in the context of a decision table. A management decision is made in the absence of knowing the true state of nature (i.e. which scenario is most correct), subsequently an important part of the process of making a management decision is having some indication of the consequences of making a decision if the true state of nature is different from that used to make the decision. Often a decision is made that performs ‘best’ (perhaps in terms of future stock sizes and removals) over the different states of nature without necessarily being the ideal management response for any single state of nature.

Table 7: Template for the projections results for three assessments (blacknose shark, dusky shark, and sandbar shark).

Blacknose Shark (Atlantic)

	Terminal conditions					$F_{Y\ rebuild}$		$TAC_{Y\ rebuild}$		P(rebuild by 2027)		
Scenario	F_{2009}	F_{2009}/F_{MSY}	S_{2009}/S_{MSY}	$Y_{F=0(P=0.7)}$	$Y_{rebuild}$	P=0.5	P=0.7	P=0.5	P=0.7	F_{2009}	TAC_{2009}	$Y_{rebuild(P=0.7)}$ TAC_{2009}
RW1												
..												
..												
..												
RWx												

Dusky Shark

	Terminal conditions					$F_{Y\ rebuild}$		$TAC_{Y\ rebuild}$		P(rebuild by 2408)	
Scenario	F_{2009}	F_{2009}/F_{MSY}	S_{2009}/S_{MSY}	$Y_{F=0(P=0.7)}$	$Y_{rebuild}$	P=0.5	P=0.7	P=0.5	P=0.7	F_{2009}	TAC_{2009}
RW1:Base	0.056	1.59	0.44	2059	2099	0.026	0.021				
RW2:High M	0.034	2.01	0.42	2150	2190	0.010	0.005				
RW3: U-shape M	0.026	1.39	0.41	2107	2147	0.009	0.005				
RW4:Hi Prod	0.080	1.49	0.45	2041	2081	0.046	0.042				
RW5:Low Prod	0.030	4.35	0.50	2217	2257	0.003	0.001				

Sandbar Shark

	Terminal conditions					$F_{Y\ rebuild}$		$TAC_{Y\ rebuild}$		P(rebuild by 2070)	
Scenario	F_{2009}	F_{2009}/F_{MSY}	S_{2009}/S_{MSY}	$Y_{F=0(P=0.7)}$	$Y_{rebuild}$	P=0.5	P=0.7	P=0.5	P=0.7	F_{2009}	TAC_{2009}
RW1											
..											
..											
..											
RWx											

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

There are many types of uncertainty that can be considered when undertaking a stock assessment. The first is the uncertainty within a single model which is typically estimated with some statistical procedure; a second is the uncertainty across different structural models (e.g. models with either different model structure or alternative assumptions about data inputs or biological assumptions); and a third is the impact of new data on parameter estimates, an approach known as 'retrospective analysis' which can be useful for determining the potential for bias in parameter estimates. Often it is found that the structural uncertainty is greater than the within model uncertainty for a range of plausible structural models. All three approaches were applied to these assessments.

The RP concluded that the AT has used and applied appropriate methods to characterize uncertainty in the four stock assessments. Approximated normal standard errors and likelihood profiles were used to characterize the uncertainty in both model parameters and other model outputs of interest, e.g. stock status in relation to benchmarks. Extensive sensitivity analyses were used to characterize structural uncertainty and retrospective analysis was used to assess the change in parameter estimates as new data was added.

The importance of structural uncertainty was recognized by both the AT and RP. Because it was not considered appropriate to describe stock status with a single model, estimates of uncertainty in model parameters and key model outputs are provided under ToRs 3 - 5 where multiple model runs are used to characterize the status of the stock and the expected response to future management through projections.

Future assessments could consider additional approaches to characterize uncertainty. Within model uncertainty could be characterized using full Bayesian integration and this could also provide insights into model fit / convergence. Most of the sensitivity analyses were a single change from the reference model in either some model assumption or data input. Some of these changes represented plausible alternative states of nature and some changes were independent of others. In such circumstances it can be useful to evaluate all possible combinations of the sensitivity analyses, e.g. make several changes at the same time. Through the automation of model running procedures such extensive sensitivity analyses can easily be implemented.

Notwithstanding this, the RP reiterates that the approaches used by the AT are appropriate and the RP have used the information provided by the AT to characterize the bounds of uncertainty in stock status under ToRs 3-5.

ToR 7: Ensure that stock assessment results are clearly and accurately presented in the Stock Assessment Report and that reported results are consistent with RP recommendations. If there are differences between the assessment workshop and review workshop due to reviewer's requests for changes and/or additional model runs, etc., describe those reasons and results.

Assessment documentation prepared by the AT was comprehensive for use of input data, model description, results and sensitivity analysis selection. An area that requires considerable improvement is in the provision of model diagnostics. Evidence of convergence should be included particularly for the base case, minimally in the form of convergence statistics (such as Geweke convergence diagnostic, Gelman and Rubin diagnostic and the Heidelberger and Welch stationarity and half-width tests for key model parameters) and preferably as MCMC diagnostic plots. Sensitivity results should include for the base case and all sensitivities, as columns, the individual objective function components (in this case, each abundance index, catch or effort series, parameter priors). Developing national standards in stock assessment documentation (e.g. Crab Plan Team 2009) should be implemented.

ToR 8. Evaluate the SEDAR Process as applied to the reviewed assessments and identify any Terms of Reference that were inadequately addressed by the data or assessment workshops.

In spite of some of the uncertainties in the conclusions, the RP believes this SEDAR process has, overall, led to a comprehensive assessment of these stocks. The DW and AW reports summarize a considerable amount of information, and providing the background documentation affords the opportunity for a thorough review. In this assessment, an independent reviewer participated in the data workshop and the AW report was also reviewed prior to being finalized for the RW. After discussion, it remained unclear to the RP whether placing greater emphasis on reviews earlier in assessment processes will automatically lead to better assessments. The utility of this approach is likely case specific depending on issues encountered during the process.

Notwithstanding the findings of this review, as well as the comments and recommendations in this report, the RP believes that the ToRs of the data and AW were generally met. For example, although the AT did not provide maps showing the geographic distributions of the individual stocks, but the descriptions in the text delineated the stocks sufficiently. The data workshop and AW reports provide detailed summaries of the information available for these species.

There is recommendation in the AW report that more time should be available for the data vetting process, while at the RW, the time available for the assessment modeling appeared to be one of the factors limiting further model development. The review of four stocks in four days at the RW was only possible because three of the stocks used the same model and limited time was placed on the review of the data inputs. The RP endorses the AT recommendation that no more than probably two stocks be assessed at one time with the same number of participants. The RP notes that the time required for a RW depends not only on the number of stocks, but also on the complexity of the individual assessments.

ToR 9. Consider the research recommendations provided by the data and assessment workshops and make any additional recommendations or prioritizations warranted. Clearly denote research and monitoring needs that could improve the reliability of future assessments. Recommend an appropriate interval for the next assessment, and whether a benchmark or update assessment is warranted.

The AT provided several research recommendations in the data workshop and AW reports, and these are endorsed by the RP to the extent that they will improve the assessment. The RP considers research leading to an improved understanding of landings and removals, that improves consistency among indices, that reduces variability within the individual indices, and that leads to development or application of a model that more fully takes advantage of the length and age data including integration of the selectivity estimation into the assessment to be priorities.

With respect to further life history research, the RP considers the following to be priorities:

- Research on post-release survival by fishing sector and gear type should lead to improved landings and removals time series

- Research on fecundity and reproductive frequency should lead to an improved understanding of population productivity. As shown in assessment, status with respect to benchmarks is relatively robust to assumptions about overall productivity; however abundance and fishing mortality rate estimates are sensitive to this information. Research about natural mortality would also lead to a better understanding of productivity but traditionally has been difficult for most species.

- As noted throughout this report, the lack of age data was a limiting factor in this assessment and collection of sex-specific age and length data would aid the assessment. Regular collection of age data will help in the construction of improved age-length keys, in the interpretation of indices particularly in cases where populations have spatially structured with respect to age, and significantly aid in fitting the selectivity within the models. Additionally, if the abundance indices are age-structured, population responses to management actions should be detectable earlier than if the indices only provide information on total abundance.

- Although information about stock structure is important, as noted under ToR 1, genetic studies may not necessarily be informative about structure. Tagging studies to determine stock structure need to take into account that populations may be discrete during reproduction, but otherwise mixed most of the time. Increased international collaboration (e.g. Mexico) could help ensure wider distribution and returns of tags.

With respect to the abundance indices, the RP recommends:

- Evaluation of the individual indices via power analyses to determine whether they are informative about abundance trends. The majority of indices used in these assessments exhibited greater inter-annual variability than would be expected given the life history of these species, and given this variability, may only be able to detect large changes in abundance which are not expected to occur rapidly. A power analysis would help to determine how much abundance would have to change in order for the change to be detected with the survey, and additionally, if the survey effort needs to be increased or re-distributed in order to be able to evaluate the effectiveness of rebuilding strategies given the relatively low population grow rates for these species.

- A small study on how to make the best use of the knowledge of the data workshop participants for developing index rankings.

- Ensuring that, to the extent possible, information about sex, length and age is collected for the reasons provided above.

With respect to the landings and removals, the RP recommends:

- Research that improves the understanding of historical landings, both in the modern and historical period and to support the assumptions about when stocks are at virgin biomass if this assumption is carried forward in future assessments. This is particularly important for GoM blacknose sharks given the difficulties reconciling the abundance indices, landings and life history information.

- As recommended by the AT, improved observer coverage particularly during periods of regulatory or gear changes (e.g. TEDs).

- Ensuring that, to the extent possible, information about sex, length and age is collected for each fishery in order that selectivity can be estimated in the model.

With respect to the assessment models, the RP recommends further model development using both simpler and more complex models taking the following into consideration:

- The RP noted that the models used in this assessment were reasonably suited to shark life history. However, other models (e.g. SS3) could also be adapted. If reproduction is modeled as a function of the number of mature females, uncertainty in the reproductive frequency, fecundity and pup-survival can be integrated into a single parameter (the slope at the origin of the SR function), and information about these traits can be incorporated via priors on the parameter. The RP recommends consideration of this approach if information on reproduction remains uncertain.

- Estimating the fishery and survey selectivities within the assessment model.

- Development of a two sex model for more direct estimation of the spawning stock

- Fitting the model to either length or age data. In addition to being necessary in order to estimate selectivities, these data can be informative about changes in age-specific abundance.

- Exploration of models that do not require an assumption that the population is at virgin levels at some point in time.

- If external age-length keys are used in future assessments, development of a key based on a growth model to better assign proportions-at-age in each length class.

- Simulation tests (management strategy evaluation) can be used to test the performance of alternative assessment methods (including the catch-free model, ASPM, ASPIC, SS3, or stock specific models), recruitment parameterizations, harvest control rules, assessment frequency and data collection. Simulation studies may have a particular use in these assessments because of the particular biology of sharks and the data poor nature of these stocks.

In the case of GoM blacknose shark, the appropriate interval of the next assessment depends on progress made towards reconciling the issues raised during this assessment process. For Atlantic blacknose shark, dusky shark and sandbar shark, the RP recognizes that population growth is expected to be relatively slow, but that modifications to the model may result in a different assessment of status. Benchmark assessments are recommended once the modifications are made. Additionally, for dusky shark, given the retrospective patterns in the present analysis and the resulting uncertainties in the assessment, updates using the existing model in the shorter term are also recommended. In the longer-term, development of a set of indicators (age-structure, total mortality estimates from catch curves, changes in abundance indices values) that could be used to determine whether status has changed sufficiently to warrant a full assessment, is recommended.

ToR10. Prepare a Peer Review Summary summarizing the Panel's evaluation of the stock assessment and addressing each Term of Reference. Provide a list of tasks that were not completed, who is responsible for completing each task, and when each task will be completed. Complete and submit the Final Summary Report within 3 weeks of workshop conclusion.

This report documents the Panel's findings with respect to each Term of Reference. Each section of the RP's report was assigned to a panelist for drafting, and was discussed during the RW. The plan developed for completion of the report is as follows: the reviewers provide draft sections of the report to the Chair by April 28th, the Chair compiles the sections and produces a complete draft of the RW report and returns it to the Reviewers by May 3rd, the reviewers provide edits, additions, clarification and other comments back to the Chair by May 6th, the Chair compiles incorporates these changes into the report and returns it to the Reviewers for final review by May 10th, and the reviewers approve the final report by May 12th.

References

Brooks, E.N., Powers, J.E. and Cortés, E. 2010. Analytical reference points for age-structured models: application to data-poor fisheries. ICES Journal of Marine Science, 67: 165-175.

Crab Plan Team 2009. Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Ateutian Regions. North Pacific Fishery Management Council, Anchorage, AK.

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 Technical Memo. NMFS-F/SPO-31, 54p. National Technical information Center, 5825 Port Royal Road, Springfield, VA 22161.

Appendix I

SEDAR 21 HMS Sharks
Review Workshop Participants

Workshop Panel

Larry Massey, Chair.....NMFS SEFSC
Jamie Gibson..... CIE Reviewer
Neil Klaer..... CIE Reviewer
Shelton Harley ..... CIE Reviewer

Analytic Representation

Enric Cortés ..... NMFS SEFSC Panama City
Kate Andrews.....NMFS SEFSC Beaufort
Paul Conn..... NMFS AFSC

Rapporteur

Ivy Baremore ..... NMFS SEFSC Panama City

HMS Representation

Karyl Brewster-Geisz..... NMFS

Observers

..... SERO
..... SERO

Staff

Julie Neer ..... SEDAR
Tyree Davis.....NMFS Miami