APPENDIX B TABLE OF CONTENTS

Appendix B Table of Contents.. B-i

B.1 Report to Directed Shark Fisheries, Inc. on the 2006 SEDAR 11 Assessment for Sandbar Shark..B-1

B.2 National Marine Fisheries Service's Response to: Report to Directed Shark Fisheries, Inc. on the 2006 SEDAR 11 Assessment for Sandbar Shark.........B-17

B.3 Scientific Review of Hester and Maunder's "Report to Directed Shark Fisheries, Inc. on the 2006 SEDAR 11 Assessment for Sandbar Shark" (2007)...........B-18

Report to Directed Shark Fisheries, Inc. on the 2006 SEDAR 11 Assessment for Sandbar Shark

Prepared by

Frank Jay Hester, PhD
Directed Shark Fisheries, Inc.

Mark Maunder, PhD
Inter-American Tropical Tuna Commission

Abstract

The Review Panel for SEDAR 11 (Large Coastal Sharks) was held 5–9 June 2006 at Panama City, FL. The panel was conducted by the Center for Independent Experts (CIE) and concluded that:

"The population model and resulting population estimates were the best possible given the data available.

"Stock status was determined from the results of a range of model fits reflecting the Panel's uncertainty about life history parameters. All results indicate that the stock is overfished and that overfishing is occurring. The target year to rebuild the stock is estimated to be 2070."

At the request of Directed Shark Fisheries, Inc., we reviewed the data and modeling of SEDAR 11 upon which the CIE based their conclusions. After review, we concluded that:

- ***The assessment proceeded without using the largest data set available, the BLOP data, which inter alia shows that average age of the catch has not declined over time, as it should if the stock were being overfished.***
- ***The BLOP data also show that the selectivity curve used for the commercial catch is wrong and needs to be re-examined.***
- ***Catch-rates for recent years remain level indicating a population in equilibrium; overfishing is not occurring, whereas the model trajectory indicates a continuing decline in abundance.***
- ***The assessment used several catch-rate series (LPS and NMFS – NE) that were either inappropriate, or did not include the available (but withheld) size and sex data (VA LL).***
- ***The age-at-maturity ogive was derived from a study that is technically flawed.***
- ***The biological parameters used in the model were selected subjectively and there may be some evidence that different values are more appropriate.***

If NMFS relies on this technically flawed assessment to make the formal finding that the stock is overfished and overfishing is occurring a legal process will begin that will require a severe reduction in TAC equivalent to closing the fishery. There is time yet to revisit the assessment before that reduction is in place if NMFS is willing to devote the effort to address the concerns that the CIE raises in their Report and we have raised in ours. Redoing the maturity ogive study may not fit into this period, but the other corrective work could be done a matter of months.

Background.

The Review Panel for SEDAR 11 (**S**outh **E**ast **D**ata and **A**ssessment **R**eview) of Large Coastal Sharks was held 5–9 June 2006 at Panama City, FL. The panel was conducted by the Center for Independent Experts (CIE) and concluded that:

- *The population model and resulting population estimates were the best possible given the data available.*
- *The change in stock status in the 2006 assessment from the more optimistic status in 2002 appears to be mainly attributable to revisions to the life history parameters in the current assessment. The population is assessed to be less productive than was assumed in 2002.*
- *In 2006, the 3-part SEDAR process of data workshop, assessment workshop, and review workshop was adopted for large coastal sharks. This process resulted in a more thorough review at all stages of the process, which was not possible with the previous stock assessments. For this reason and those concerning the life history parameters given above, the Panel is confident that the 2006 assessment gives a more reliable estimate of stock status than obtained from the 2002 and earlier assessments.*
- *Stock status was determined from the results of a range of model fits reflecting the Panel's uncertainty about life history parameters. All results indicate that the stock is overfished and that overfishing is occurring. The target year to rebuild the stock is estimated to be 2070.*

Directed Shark Fisheries, Inc. (DSF) represents several entities involved with the commercial fishery for Atlantic large coastal sharks. The group disagrees with these conclusions, which are at variance with their observations of the fishery. There is no indication of a continuous decline in either catch rate or fish size (average carcass weight) predicted by the modeling. Of particular concern to the fishermen is the determination that the fishery for sandbar shark needs to be closed for a 65-year rebuilding period. Directed Shark Fisheries, Inc. has asked us to review the data and modeling of SEDAR 11 upon which the CIE based their conclusions to attempt to reconcile the two different perceptions of the status of the sandbar stock, report our findings and make such recommendations as may be appropriate.

The Problem

The CIE in reaching their conclusions stopped short of taking the vital but simple step of comparing the model results with actual information from the fishery. A cursory examination would show that the commercial landings and catch rates have remained stable for over a decade, and catch-rate (abundance) indices are mostly flat or trend upward over this period. These observations are inconsistent with the model output, which indicates a steady decline in biomass over the same period. The problem this created is that the CIE and SEDAR are pronounced by NMFS to have provided a peer review approval of this assessment, "…the best possible given the data available." The fishery now likely faces a major reduction in TAC under current law.

The CIE accepted both the data and analyses provided by SEDAR 11 and the conclusion that the stock is overfished and that overfishing is occurring with some caveats, and raised a number of issues for future examination. The issues raised by the CIE are important; so important in fact; that we wonder why the CIE did not express greater concern about the confidence that can be put on the SEDAR 11 assessment and recommend that some issues be addressed before the assessment was accepted.

These concerns might have been more strongly phrased had the CIE been advised that some of the data they highlighted for future work were actually available but not presented at SEDAR 11. We will now make use of additional data to explain some of the inconsistencies between the perception created by the model results and the perception held by the commercial sector.

The Data

The Review Panel qualified their conclusions by stating:

Research recommendations are included in the reports from the Data and Assessment Workshops (and in 2.3 below), so what follows is not intended to replace them but rather to emphasize specific needs for sandbar shark.

Two recommendations in particular are extremely important. These are:

Issue: A number of catch-rate indices were used, and it was not obvious which components of the sandbar population they were monitoring.
• Using information on the size composition of catches from these indices, if available, would be helpful
• Maps of where (and when) the catch-rate series are located, along with the location of the fisheries, would aid in interpreting these series

Issue: The assessment used an age-structured model, but no age information was used.
• The predicted age compositions for the population and the catch in the model may provide useful diagnostics for the performance of the model. Research should be directed into developing these diagnostics, including verification with any available data on age composition. One example of a diagnostic indicator is the mean size/age in the catch and population, and from any catch-rate index that may collect size composition data…

Size, sex, location and other information are contained in two data sets used at SEDAR 11 and this additional information was available to SEDAR 11 and the CIE, but was not presented. One set is the Bottom Longline Observer Program (BLOP)[1], the other the VIMS longline survey (VA LL). The BLOP comprises observed sets during the period from 1994 through 2004 from N. Carolina south and into the eastern Gulf of Mexico and covers all seasons and most of the range of the commercial fishery using a gear (bottom longline) that accounts for nearly 90% of the commercial landings. The latter, the VA LL, comprises sets from an intermittent summer longline survey from 1974 through 2004 confined to a small area off Virginia.

The BLOP data for 1248 observed sets were used to develop a catch-rate index at SEDAR 11, but the size and sex composition of the catch was not made available at SEDAR 11. The VA LL data were presented to SEDAR 11 in summarized form with no detailed information. NMFS standardized the series after the Data Workshop ended using the limited data provided that did not include age, size or sex. Through the cooperation of NMFS and University of Florida, we were provided with extracted BLOP data that includes length and sex and reproductive state information, general location (we were not given precise locations for the sets because of confidentiality concerns) and some environmental information. For the VA LL series, we do not have the data set available to SEDAR 11. The Principal Investigator for the VA LL survey declined to provide age, size or sex information until he has analyzed and published his 30-years of data.

The BLOP data set is useful for several reasons:
1. It is arguably representative of about 90% of the commercial catch of sandbar (but see bullet 3).
2. It provides length and sex information on all sandbar taken including discards (which were few) and should be a reasonably unbiased sample of the commercial catch.
3. It covers the South Atlantic Region and the eastern Gulf of Mexico Region (where most of the sandbar catch occurs). This is most of the range of the fishery. It does not include the North Atlantic Region.
4. It covers all months when fishing is allowed.

The VA LL data set includes information from 637 bottom longline sets beginning in 1973 and running through 2004. No sets were made in some years. The number of sets in any year varied from 3 to 47. There were 371 sets made between 1995 (none in 1994) and 2004, the same period covered by the BLOP data; however, the two areas do not overlap.
The standardized index used in the assessment was done after the Data Workshop and the procedure omitted the years prior to 1981. The index is not size or age specific, but assumes that the selectivity curve used for the commercial fishery should apply.

In addition to the above data sets, we received a copy of the State-Space Age-Structured Production Model (SPASM) from Dr. Liz Brooks, NMFS, and we will refer to several SEDAR 11 documents.

Analyses and Results

1. **BLOP Data**

[1] A.k.a. PLLOP and Commercial Shark Fishery Observer Program (CSFOP).

Length frequency samples.

The BLOP data set contains length measurements on 21,031 individual sandbar sharks. The distribution of the sample lengths by sex is shown below (Fig.1).

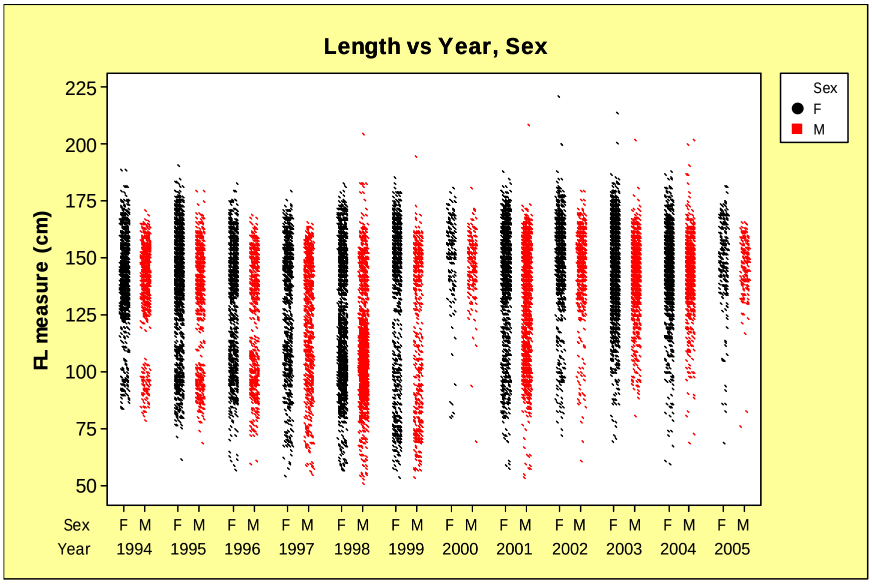


Figure 1. Lengths of individual sandbar sharks taken during Bottom Longline Observer Program trips

These 21,031 length frequency samples are important for two reasons. The allowed us to look for changes in the size (age) composition of the population over 12-years of exploitation, and they provide an indication of the pattern of selectivity of the bottom longline gear.

Change in age composition.

The average age (size) in a population of fish under exploitation is expected to decrease. This is particularly true for populations of long lived fish like sandbar.

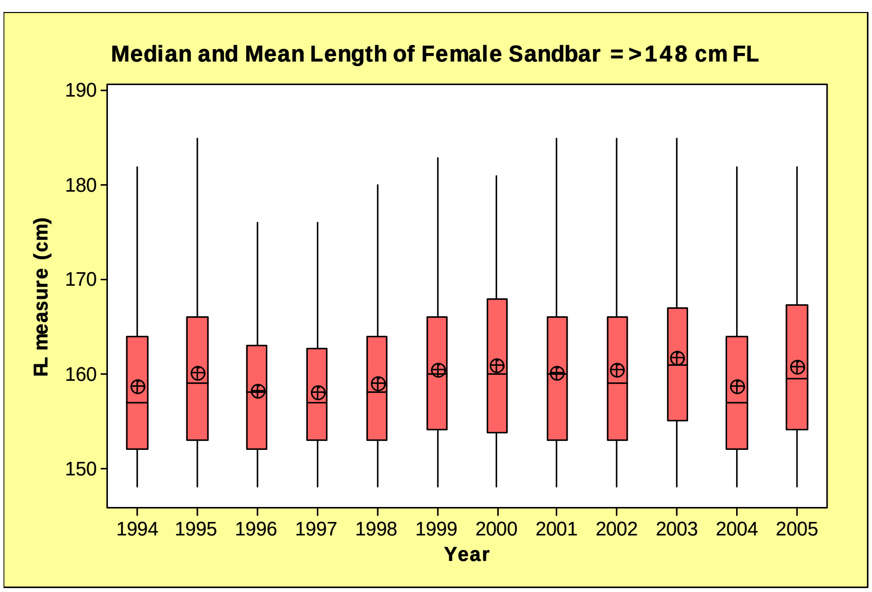


Figure 2. Median length mature females; means are indicated by circled cross symbol

The SPASM Model predicts a 45-percent decline in spawning biomass over these years, which should be reflected by a decrease in the average age of spawners (taken as >148 cm Fork Length). The BLOP data on the other hand indicate a stable size or slight increase in average size (Fig. 2) over the period. How this should be interpreted is arguable, but if size at age is constant as the model assumes the observed data are at variance with the model prediction.

<u>Selectivity.</u>

Converting lengths to ages using a von Bertalanffy equation (Sminkey and Musick, 1995) gives the distribution for the BLOP catches shown in Figure 3.

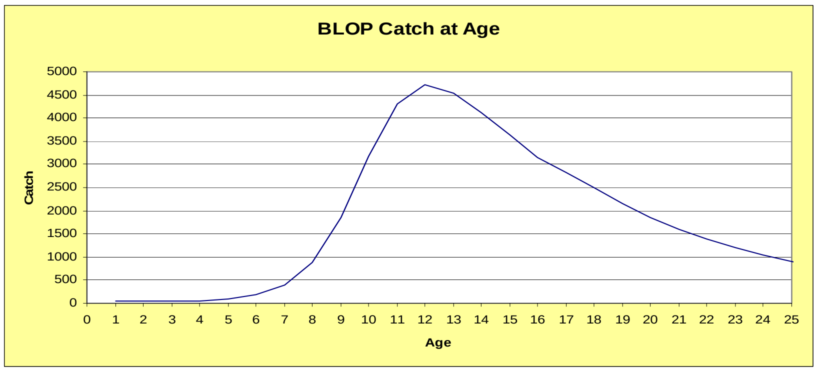
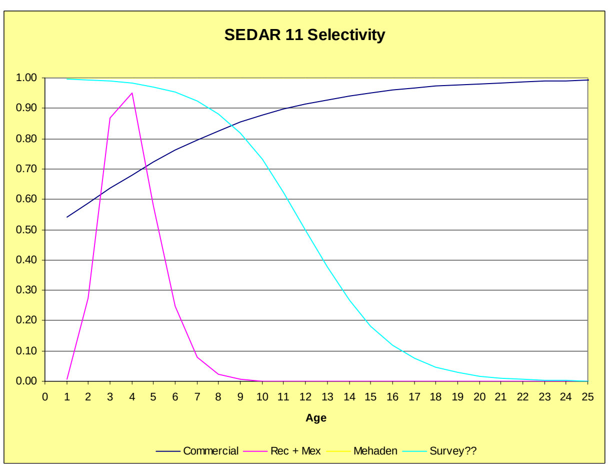


Figure 3. Sandbar catch at age from the observed "commercial" catch.

SEDAR 11 developed a series of curves believed representative of the selectivities in various sectors (fleets) of the fishery. These are reproduced in Figure 4a below. A revised selectivity curve is shown in Figure 4b.



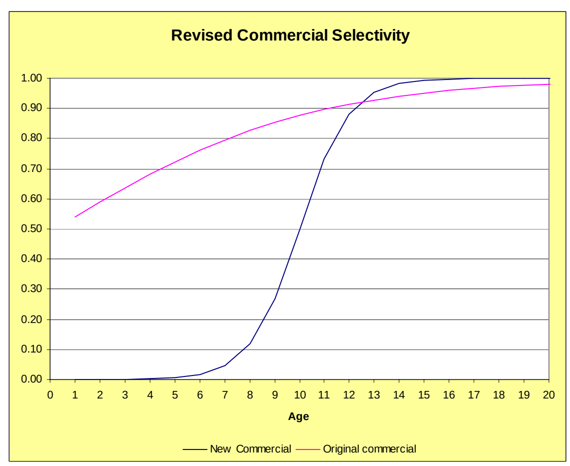


Figure 4a. **The four selectivity curves used for the 2006 assessment.**

Figure 4b. The SEDAR 11 commercial selectivity curve and the BLOP based commercial curve

The plot indicates that juvenile sandbar sharks are less vulnerable to the commercial gear than was assumed by the SEDAR 11workshop.

<u>Time and area differences.</u>

The catch-at-age and selectivity patterns estimated in Figures 3 and 4b were derived using all BLOP observations combined. For the BLOP program, fishing takes place in three Regions (not the same as the three Regions used by the HMS management plan). The BLOP Mid-Atlantic Bight Region does not extend north into Virginia and there were few sets made north of 37° N. The HMS North Atlantic Region begins off

Virginia, so that the HMS South Atlantic Region comprises both the BLOP Mid-Atlantic Bight Region and the BLOP South Atlantic Region.

The BLOP data set include information by region and date. There are significant differences in average size among regions and seasons. These are shown in Figure 5, suggesting that a single selectivity may not be appropriate for all regions and seasons. In particular, season one in the mid Atlantic bight catches smaller individuals. This region is closest to the area used for the VA LL survey and indicates that the selectivity for the VA LL survey may also be different from the commercial selectivity used in the model.

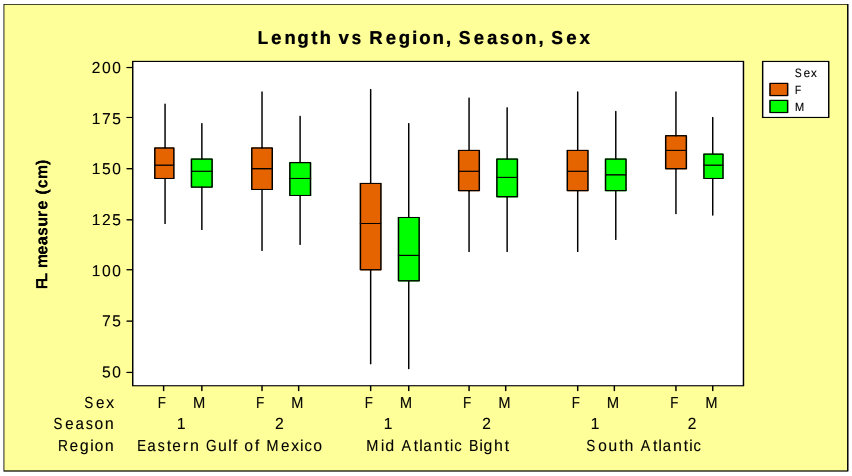


Figure 5. Median size by BLOP Region, Season 1 = Jan-Jun, 2 = July-Dec, and sex

<u>Sex ratio differences.</u>

Figure 6 indicates that bottom longline gear is selective of females. The overall ratio from the BLOP is 1:1.31 male to female. Whether this reflects a true sex ratio difference in the population or a targeting and/or segregation by sex deserves further investigation.

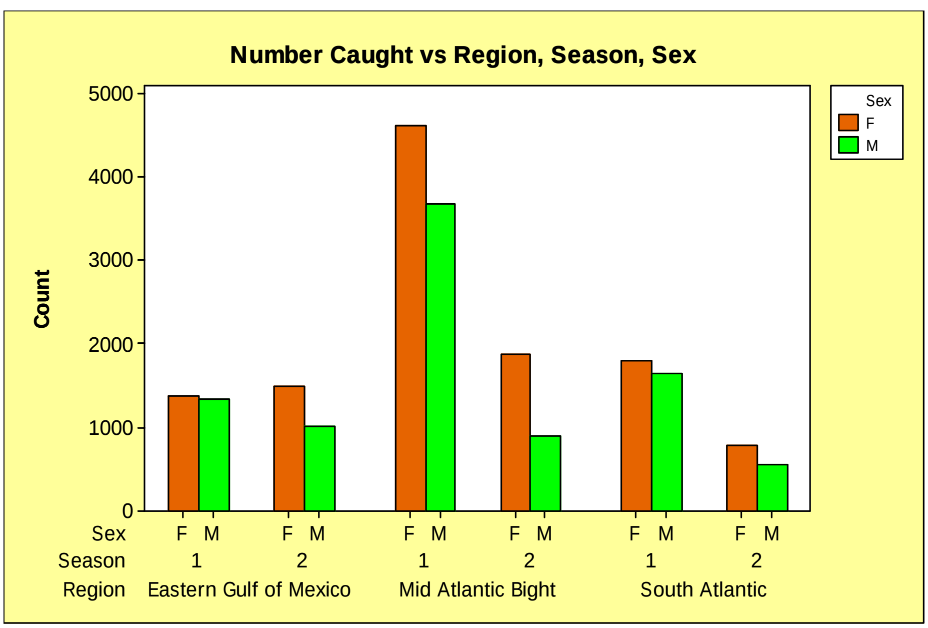


Figure 6. BLOP catches showing a preponderance of females, especially in the Mid-Atlantic Bight

2. SPASM Model

In this section, we will look at the modeling and consider how some changes in the inputs effect the perception of the status of the sandbar stock, as well as look briefly at the model itself. The model of interest for the sandbar assessment is Shark_SPASM.

"This model is [with some modification] the model used in the 2002 Large Coastal Shark was a state space, age structured production model (SSASPM, Porch 2002). Unlike a production model, the SSASPM can incorporate age-specific differences in model parameters such as growth, fecundity, and gear vulnerability (selectivity). In the case of long-lived, late-maturing fish or when there are multiple fisheries that exploit different age classes, having the flexibility to incorporate age-specific information could lead to a better fit to observed data. Age specific vectors for fecundity, maturity, and selectivity are specified by the user, and length and weight at age are calculated within the model based on user-specified growth functions. Natural mortality at age and a stock recruitment function are additional model parameters. The stock recruit function is parameterized in terms of virgin recruitment (R0) and pup survival. To derive the initial age structure for the first year that data is available, the model estimates a level of historic fishing (Fhist) and calculates the corresponding equilibrium population age structure. A historic selectivity vector is specified by the user, which is multiplied by Fhist to arrive at the historic age-specific fishing mortality rate. A historic selectivity vector of 1 for all ages was assumed.

"Continuity Model Inputs

"*Data*
"Data inputted to the model included maturity at age, fecundity at age (pups per mature female), spawning season, catches, indices, and selectivity functions …. Catches were made by the commercial sector, the recreational sector, and the Mexican fishery. In addition, unreported commercial catches were estimated, as were menhaden discards. Because of similar selectivity functions, the commercial and unreported catches were combined, and recreational catches were combined with Mexican catches, yielding a model with 3 distinct "fleets". A total of 13 indices were made available after the data workshop. The "DEL age 0" index was not used, as this model began with age class 1, which means that the stock recruitment relationship governed the number of one year olds to survive from the initial number of pups produced in a given year. Catch data begin in 1981, while the earliest data for the indices is 1975 (VA-LL). The missing catch for years 1975-1980 was treated several ways: the model estimated the missing catch; the missing catch was filled in with either the series-specific average, or series-specific assumptions were made….

"*Parameters*

"Estimated model parameters were pup survival, natural mortality (ages 1+), virgin recruitment (R0), catchabilities associated with catches and indices, and fleet-specific effort. In some models, a level of historic fishing (F_{hist}) was estimated, while other models fixed this parameter at 0 (assumes virgin conditions in 1975)."

(Quoted from SEDAR11-AW-03)

We investigated the sensitivity of the stock assessment model's results to assumptions about 1) the catch data, 2) the indexes, and 3) the mechanics of the model with the assumed biological parameter such as natural mortality (M) and fecundity,. The catches are of two sorts, the level of historical fishing and the estimated catches for which there are data. However, it quickly became apparent that this would be too large a task for this type of report, and instead will highlight a few examples that will indicate where there appear to be problems that need to be addressed.

i. Catch data As set forth in Liz's explanation of Shark_SPASM, the catch data comes in two parts, the historic catch that the model estimates, and the recorded catches starting in 1981. The catch before 1981 was assumed while from 1981 on it was based on estimates (recreational surveys of catch) or from recorded landings (commercial). The historic catch is estimated from the model. The recorded catches are for several sectors: commercial, recreational, scientific, Mexico, menhaden by catch, and discards. Most are estimated from sample data and dealers' reports.

The commercial catches are probably as good as can be had, but the recent discovery that there was major problem with the recording of the dealers' landing reports may result in some modifications of the estimates for the past few years.

There is no reason at this time to expect that adjustments can be made to catches for the other five sectors. The recreational catches are known to be highly uncertain, and should be subjected to more extensive sensitivity runs than has been the case. The other catch estimates are relatively minor in numbers and any changes unlikely to have any significant effect on the assessment.

Although recorded catches are assumed to begin in 1981, the model base case result (Fig. 7) assumes the stock biomass was virgin in 1975 because the first year for which there was an indexing value (the Virginia Longline or VA LL) was said to be 1975. However, when the VA LL index was standardized for SEDAR 11 to use in the model it was found that the earlier years lacked the information needed for the standardization. This complicated the modeling, as the first year having a standardized index now was 1981and, since F_{hist} was assumed to end with 1975, some way had to be found to bridge the gap to 1981.

For modeling the stock from 1975 to 1981, catch information was used from 1975 to 1981. This was estimated assuming that the recreational catches were zero in 1975 and increased linearly from 0 in 1975 to the estimated number in 1981and that the commercial catches were as in 1981. The slow decline in SSB/B0 between 1975 and 1981 shown in Figure 7 results from the recreational catches, which are the only appreciable catches assumed. Catch is the only thing that makes this model decline as there is no annual random variation in recruitment (and no catch-at-age data to estimate it). Recreational catch is believed to target young sharks and therefore some time must elapse before the effect of taking young fish shows up in the biomass of older fish. The commercial fishery, which targets larger fish, begins in the mid-1980's and, combined with the effect of the removals of the younger fish earlier on, is followed by an immediate and more rapid decline in SSB/B0 reaching a depletion level of 0.31 in 2004.

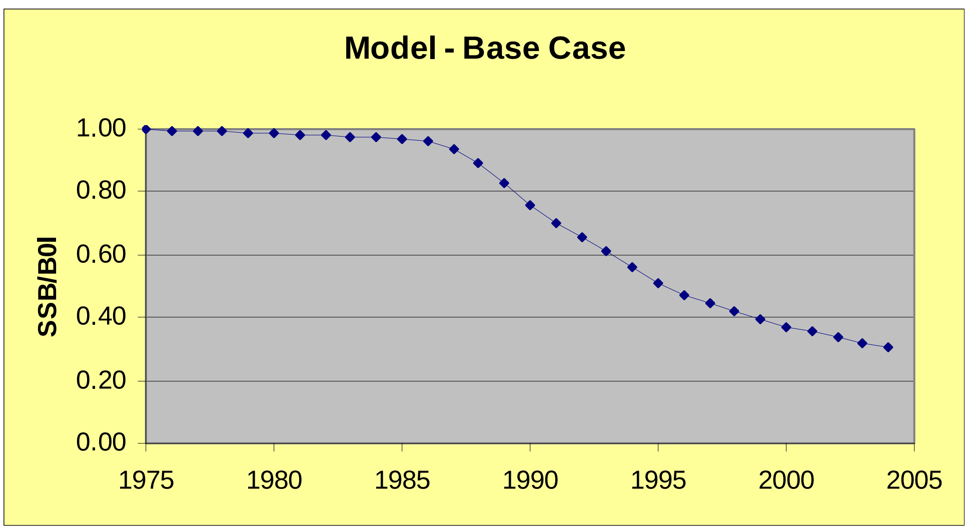


Figure 7. Trajectory for Model Base Case.

ii. Indices The eight indices used in the model are plotted in Figure 8 along with trajectory for the ratio of the Spawning Stock Biomass each year to Virgin Spawning Stock Biomass in 1975 (SSB/B0). The VA LL index, with some years missing, begins with 1981. The second longest time-series index is the Large Pelagic Survey index for recreational catch, which starts in 1985. The other indices start in 1993 when regulations for LCS first were implemented, and include indices from the commercial fishery.

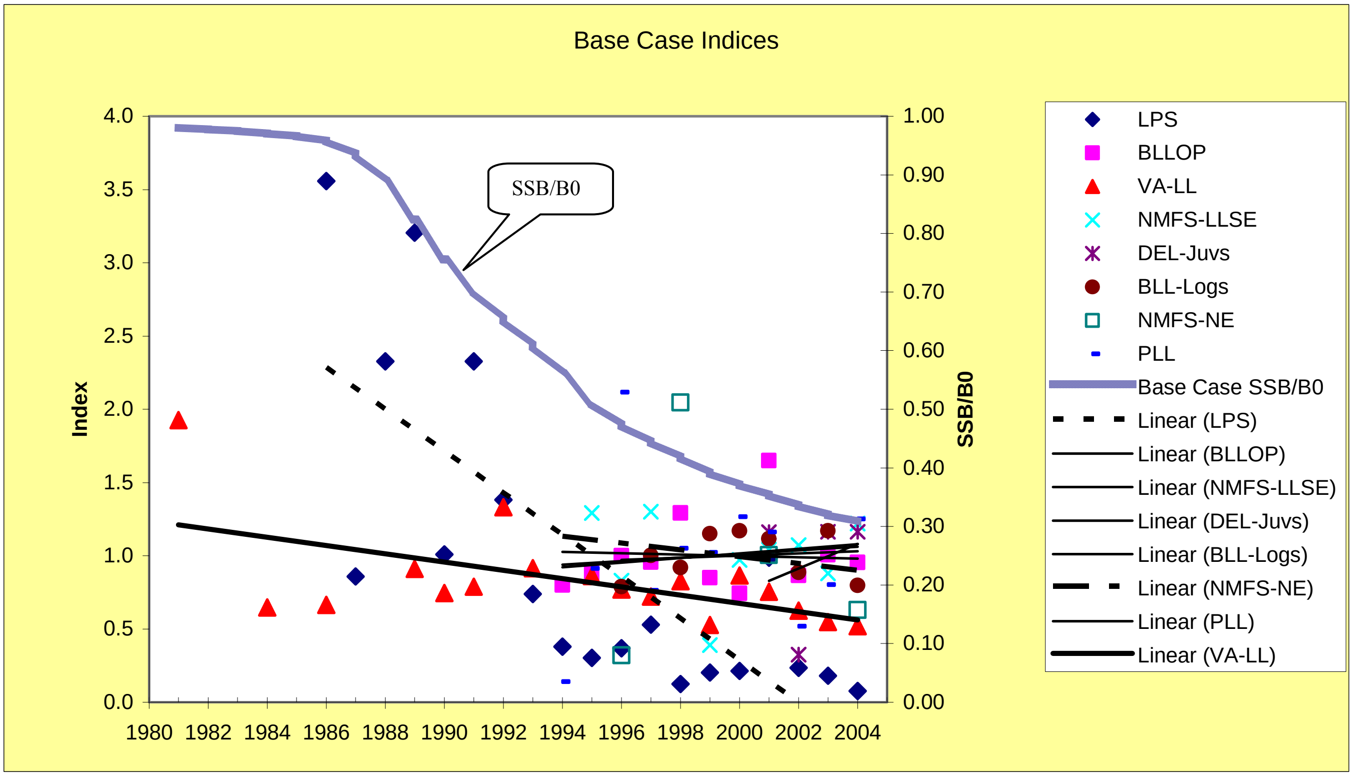
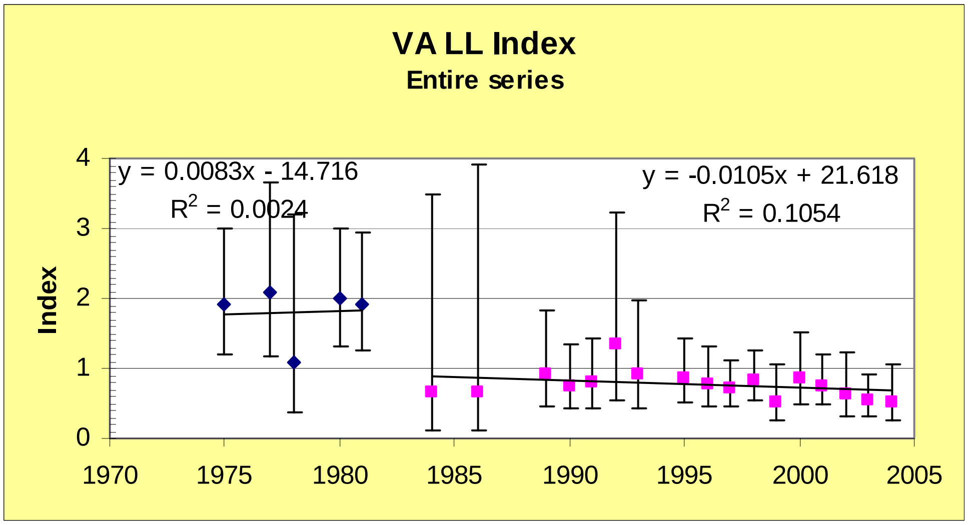


Figure 8. The eight base case indices plotted and showing their liner trends with the trajectory of the ratio of the spawning stock biomass to the virgin spawning stock biomass (SSB/B0) shown for comparison.

For the Base case, all indices are given equal weight in the analysis. This means that any index in which the points may have a trend, even though the variability (CV's or standard deviations) are very large, and the trend or slope is statistically not different from zero, will be seen by the model to be as good as an index that may have a statistically significant slope. The result is that a "bad" index (large CV) such as the LPS or NMFS NE is given equal weight to a "good" index such as the BLL-Logs.

Another problem with some indices is that they are not consistent throughout their lives. The assumption is that an index is proportional to stock abundance over time and that other factors such as fishing methods, area fished, environment, regulations, etc., remain constant or can be controlled in the course of standardizing the index. This may not be true, yet the index may be used even when some factor other than abundance is known to have changed over the course of time, as is the case with the LPS and, perhaps the VA LL.

The nominal trends for the VA LL, the LPS and the NMFS NE indices all are negative and roughly, in agreement with the biomass trajectory, which is not surprising since the trajectory is, in part, determined by the indices. Beginning with the VA LL, figures 9a and 9b show that the series consists of two parts that are essentially without a trend, an early period from 1975 through 1981, and a recent period from 1984 to present. The Index value for the early period 1975-1981 is roughly twice that for the recent period 1984-2004. The index used for SEDAR 11 omits all the years of the early part and begins with the final year 1981. Combining 1981 with the recent years causes the index o develop a negative slope that, though not statistically different from zero (flat), is perceived in the model to indicate a decline in abundance over the entire period 1981 to 2004. Why there is a difference in index level between two periods is unclear. We lack the data on size (age) and sex of the fish that might answer the question.



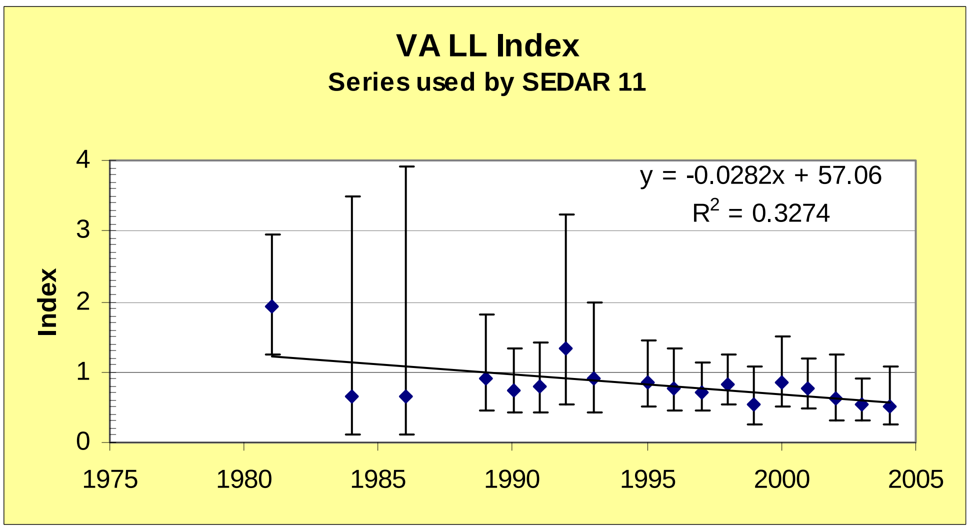


Figure 9a (top) and 9b (below). The Virginia Longline Index with trend line(s) and 95% C.I. None of the trend lines shown has a slope that differs from zero. Note that the full series beginning in 1975 appears to have an early part and a late part. The trend for each part is statistically flat. SEDAR 11, by beginning the series with 1981, produces a combined series with the 1981 point giving a larger (though still not significant) negative slope to the linear trend for the index.

The LPS index (Figure 10) has the same difficulty as the VA LL index in that it consists of two periods with high values in the early period and lower values in the recent period and addition problem that it has a very high degree of uncertainty associated with the second (recent) period. However, in the case of the LPS index we know a bit more about why the early period differs from the recent period. The LPS Index is for recreational catches off the NE Atlantic coast. The selectivity for this index was assumed the same as for the commercial catch, but no age or size information was available to confirm this supposition. This index has been used in previous assessments, but each time it was split into two indices: 1986-92 and 1993 to most recent year available. This was in recognition of the fact that the sportfishing regulations (size and bag limit) that went into effect in 1993 changed the way this fishery operated. One of us argued during SEDAR 11 DW that this should continue to be the case, or the index should not be included in the base case. That argument was dismissed out of hand. **We emphasize here that it is important to note that the LPS index is clearly two essentially flat indexes (slopes do not differ from zero), and to use the entire series to establish a trend that receives equal weighting in the assessment is not scientifically defensible.**

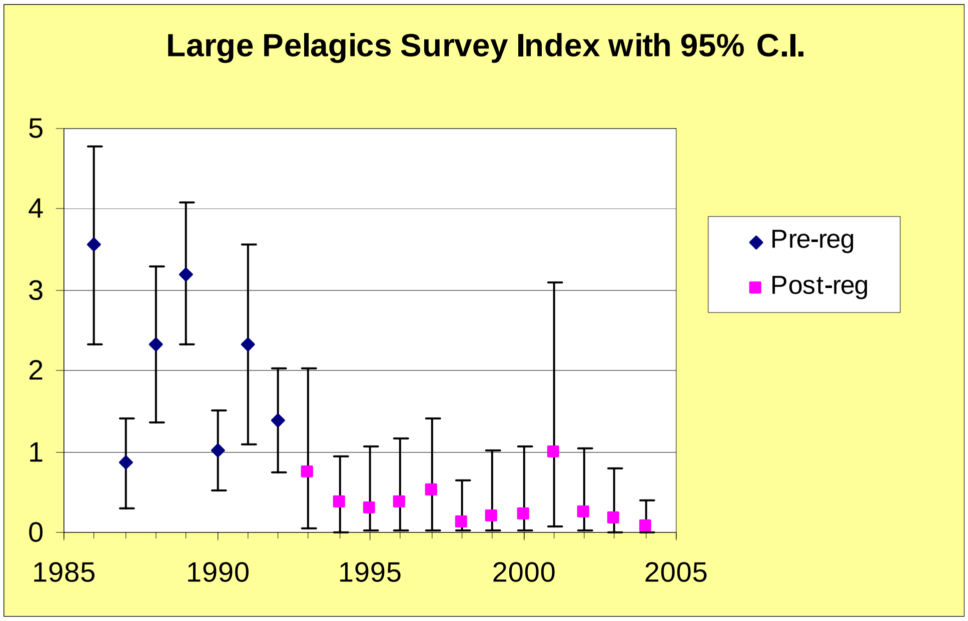


Figure 10. LPS index divided into to periods: Pre- and Post implementation of angling bag limit for Large Coastal Sharks that began in 1993. A reduced bag limit and size restrictions were added in 1999. The trend for the entire combined series is negative and significant, whereas the trends for the two separate periods are not different from zero.

The third index the NMFS-NE is a different matter (Figure 11). It is a puzzle why this index was selected as a Base Case index other than it has a negative slope when given equal weight. It has such enormous coefficients of variation that it is takes a leap of faith to accept that it contains any reliable or useful information about stock abundance. We believe that there is no valid reason to include it even as a sensitivity index.

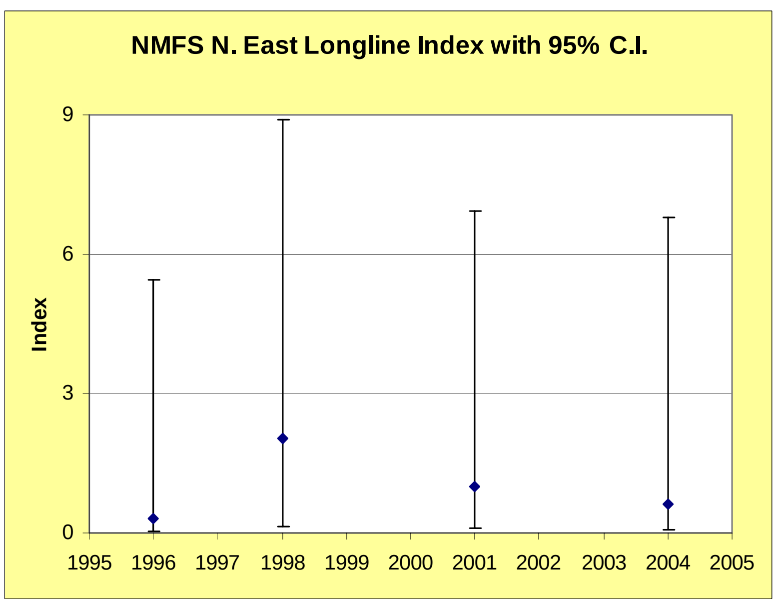


Figure 11. NMFS-NE longline index. The question is why an index with only four points and such an extreme range of uncertainty was included in the Base Case analysis. The index has no statistical trend other than zero, but when given equal weight in the assessment has considerable influence on the outcome. NMFS-NE index should not have been included as a base case index. Aside from the fact that there are only four observations, the enormous CV's should have precluded its use even as a sensitivity index.

The VA LL index may or may not be usable when and if it is properly standardized. As with the LPS and NMFS NE it samples only a fraction of the sandbar stock during the summertime when some fish have moved north to the Atlantic pupping grounds; however, large summer catches occur in waters south of Cape Hatteras and in the Gulf of Mexico at this time, thus these indices sample only a fraction of the population. None of these indices sample the areas where the majority of the fish are located and how representative these samples are of the population needs to be determined. Second, the VA LL index has in the course of sampling collected size and sex information. So far, the author has refused to make this information available. Thus, it impossible to know what size or sex selectivity to apply to the series – what segment of the population it is monitoring – a flaw with the LPS data as well. Until that information is provided, the use of this index should be restricted to a sensitivity run.

The remaining five indices are plotted in Figure 12. Three are from the commercial fishery, sample the entire range of the fishery, and begin when mandated by LCS Fishery Management Plan in 1994. What is of interest is the fact that all five indices are stable or have a positive trend over the ten-year period, whereas the model predicts the spawning stock has declined over 40-percent. **The inconsistency between the model prediction and the stable or increasing trend in abundance indicated by the five indices taken together with the failure for the average age of the catch to decline should have been a red flag to the CIE that the model has a problem that has to be corrected.**

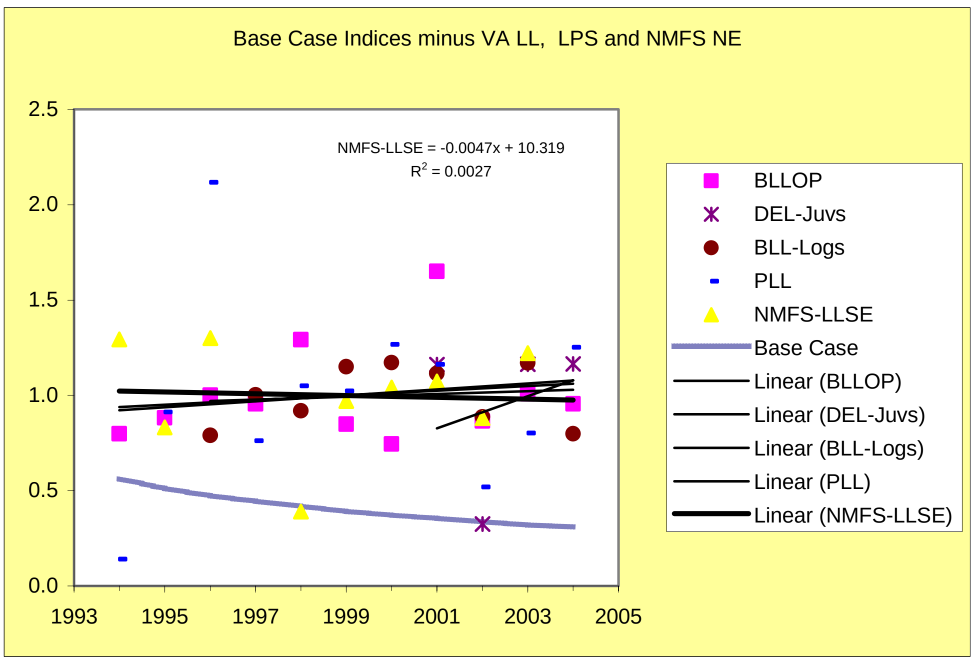


Figure 12. The five base case indices plotted and showing their liner trends with the trajectory of the ratio of the spawning stock biomass to the virgin spawning stock biomass (SSB/B0) shown for comparison. Note that one, NMFS-SE, has a very slight and non-significant negative slope.

ii (a) Testing the indices. In an assessment model, the indices establish a trend in abundance, which together with the catches and the workings in the model estimate the present condition of the stock. How much influence the indices have depends in part on the model. We tried several combinations of the indices to see how the output of the model changed depending on the combinations we selected. The different trajectories for SSB/B0 are shown in Figure 13.

ii (a) (1) Base Case and ii (a) (2) No VA LL The first trajectory to locate is the base case trajectory. If Figure 10 is not in color, the easiest way to identify the different trajectories will be to look at about the year 1995 and move up vertically. The base case is marked only by open square symbols and these are the second set of symbols from the bottom. What makes them difficult is that when we plotted the trajectory with The VA LL index heavily down weighted (the line labeled No VA LL) the trajectories are nearly identical with the base case, and the square symbols appear to be part the No VA LL curve. The final output levels for both trajectories is 31-pecent of the virgin spawning biomass This result was surprising as in past assessments the VA LL index alone had a major impact on the perception of the status of the stock. However, the re-standardization of the index done this year combined with omitting the years prior to 1981 resulted in a less steep decline than in the past. With this assessment, down weighting this index alone has essentially no effect on the model outcome.

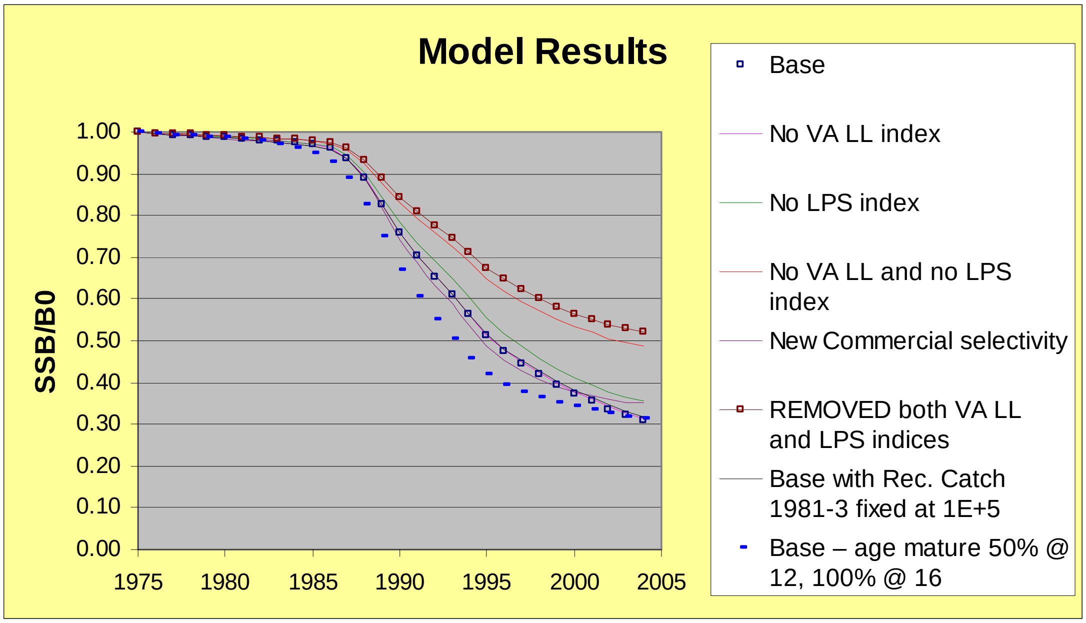


Figure 13. Experiments with the indices. See text for explanation of trajectories.

ii (a) (3) No LPS Our next experiment was to heavily down weight the LPS index. This is about equivalent to using inverse variance weighting for this index. The resulting trajectory appears as the fourth curve up from 1995 as the Base Case and No VA LL appear together as a single line. Down weighting the LPS index results in a more optimistic outcome.

ii (a) (4) No VA LL and No LPS In this experiment, both the VA LL and the LPS indices were heavily down weighted. The resulting curve is the next to the top. The result is much more optimistic, with SSB/B0 near the 50-percnet level although the stock continues to decline. Since the remaining indices are nearly flat in trend, the failure to flatten is likely driven by the biological parameters assumed in the model.

ii (a)(5) remove VA LL and LPS We then re-ran the experiment this time removing these two indices from the data file rather than merely down weighting them. The trajectory is the top curve with the closed square symbols. The final ratio is a bit above the 50-percent level and fishing mortality is less than F_{MSY} – the stock is not overfished and overfishing is not occurring. From a technical standpoint, the difference between this run and run 4 is interesting because it demonstrates that down weighting an index, which is easier to do than removing it form the data input file, is not exactly equivalent to removing it. This is mainly when the index stands alone for the first part of the time series. In the case of the VA LL and LPS they start about a decade before the other indices, therefore the small signal that remains after down weighting the index still affect the model.

ii(a)(6) Base Case using a different selectivity curve The next experiment we tried was to modify the commercial selectivity curve to be closer to what was observed in the BLOP data base. The trajectory is the second curve from the bottom. Using this selectivity curve results in a slightly more optimistic outlook, and more interesting is that here the trajectory flattens out in the recent years instead of continuing to decline as with the other runs so far. Why this happens is worth further investigation.

ii(a)(7) Base with Rec. Catch 1981-3 fixed at 1E+5 Here we chopped the early recreational catches down to a low level to see how sensitive the model is to what is a very uncertain estimate of catches. The trajectory is essentially the same as with runs 1 and 2 and overlies these two runs. This and the next run were done also by SEDAR 11.

ii(a)(8) Base case with age mature 50% @ 12, 100% @ 16 The last run in this series of experiments looks for the effect of using a young age at maturity ogive. A similar run was done by SEDAR 11and in both cases the final depletion level is the same as with the Base Case. However, we need to point out that the trajectory, which was not plotted by SEDAR 11 (this is the bottom line in the figure) shows a steeper decline than the Base Case followed by a leveling off in the last years. As the leveling off implies, fishing mortally is lower (by about half) in the terminal year than in the Base Case.

The resulting reference points for these runs are given in Table 2 along with runs 9 through 12 that are not plotted in Figure 13.

Case	SSB_{2004}/SSB_{virgin}	SSB_{2004}	F_{2004}/F_{MSY}	Pup survival	Steepness
1. Base Case – files received from Liz Brooks	0.31	428,000	3.72	0.62	0.32
2. Base down wt. VA LL	0.31	435,000	3.69	0.62	0.32
3. Base down wt. LPS	0.36	570,000	1.41	0.98	0.42
4. Base down wt. VA LL and LPS	0.49	936,000	0.98	0.97	0.42
5. Base remove both VA LL and LPS	0.52	1,080,000	0.87	0.95	0.42
6. Base with modified commercial selectivity	0.35	796,000	2.46	0.68	0.34
7. Base with Rec. Catch 1981-3 fixed at 1E+5	0.32	424,000	3.52	0.65	0.33
8. Base – age mature 50% @ 12, 100% @ 16	0.31	678,000	1.76	0.57	0.41
9. Base down wt. LPS, NMFS-NE, and VALL	0.67	1,011,000	0.92	0.97	0.42
10. As in 9 with the modified maturity	0.51	1,525,000	0.55	0.93	0.53
11. Base (1) with modified maturity and commercial selectivity	0.45	1,616,000	0.94	0.69	0.46
12. As 10 with modified commercial selectivity	0.58	2,751,000	0.48	0.85	0.51

Table 1. Reference points from experimental runs for Shark_SPASM

Runs 1 through 5 explore the effect the two long time series indices VA LL and LPS have on the model outcome. As noted above, down weighting the VA LL (2) has slight effect on the outcome whereas down weighting LPS (3) results in a more optimistic outcome with the F ratio and SSB being improved considerably, but accompanied by an estimate of pup survival that is quite high, and an increase in the estimate for steepness. Runs 4 and 5 reduce or remove the effect of both VA LL and LPS from the model and, as previously noted, provide a much more optimistic outcome, but again with a very high estimated pup survival and increased steepness.

To conclude our exploration of the negative indices, we made Run 9 that down weighted NMFS-NE as well as VA LL and LPS. As expected, there is further improvement over the optimistic outcome seem for Run 4. Again, pup survival is estimated to be quite high and steepness increases. We did not try actually removing all three indices, but we anticipate that the result would be an improvement over Run 5 with a lower F ratio and slightly lower estimated pup survival. This run has leaves the model with information mainly from indices that cover the period from 1993 through 2004 and are all essentially flat. (The down weighted indices still have a slight effect.) Thus, the outcome is the result of the catch information, the biological assumptions and the selectivity curves combined with indices that indicate stock abundance has been stable in recent years.

Run 6 investigates the Base Case using a modified commercial selectivity that is based on observational data in the BLOP data set. The outcome is similar to what we got by down weighting the LPS index: the F ratio is improved as is stock size, but with the estimate for pup survival much lower and perhaps more realistic than when LPS was down weighted.

Run 7 investigates the Base Case using a modified recreational catch that reduces the large catches in the early years to a perhaps more believable level. The outcome is slightly more optimistic, but because recreational catches are so poorly accounted, any changes to the data base are speculative. Further exploration and sensitivity runs should be done after examining the origin of the estimated catches, but that was not something we could do at this time.

Run 8 investigates the Base Case using a modified maturity ogive that is based on observational data other than the Merson study used by SEDAR 11. As noted by SEDAR 11 when it made this sensitivity run, the outcome is unchanged with respect to the final biomass ratio. However, the spawning biomass is considerably great – as might be expected since the number of mature animals would be increased by the addition of younger fish – and the F ration is much more optimistic. Pup survival and steepness are acceptable.

We then used the modified maturity ogive in Run 10 that also down weighted the negative indices. The result is very optimistic but pup survival is estimated to be high.

For Run 11 we returned to the Base Case inputs but used both the modified maturity ogive and the modified commercial selectivity. The result is optimistic with F_{2004}/F_{MSY} ratio less than 1.0 and pup survival (0.69) believable.

Run 12, the last we did, down weights the negative indices and uses the modified maturity ogive and commercial selectivity. The result is optimistic and pup survival is arguably acceptable. In this run and five other runs the stock is not overfished and/or overfishing is not occurring. In ten of the twelve cases examined, the model estimates that steepness lies outside the bounds (0.2 to 0.4) set by SEDAR 11 but there are no quantitative data to support this range. Density dependence response is presumed to exist for SB, perhaps mediated through a change in age at maturity and a lowering of natural mortality for both adults and pups, and the biological basis for fixing the upper bound for steepness at 0.4 needs to be examined.

iii. The Model

The CIE has this to say about the model: "Ultimately, the methods used for estimating stock status were found to have been much more sensitive to assumptions about life history parameters than the catch and catch-rate data used in the model.

"Size and maturity stage information was reported as being collected from the VIMS longline and some of the other series, but those data were not supplied to the stock assessment scientists. Given that the VIMS survey was a designed fishery-independent survey, it would have been helpful to have the size information to see if the component of the population that it was monitoring had been changing over time.

"An age-structured population model with state-space dynamics for some of the components and prior distributions assigned to some of the parameters was fitted to the data. No age data were used in the model, and the age structure was used mainly to incorporate different natural mortalities- and selectivities-at-age for the different fisheries (i.e. commercial, recreational, bycatch in menhaden fishery). Catch-rate indices were assumed proportional to population size, albeit with series-specific catchabilities and selection curves dependent upon whether they were commercial- or recreational-fishery-dependent, or fishery-independent series.

"The model adequately incorporated the information from the available catch-rate indices and was the best available for the data provided. However, while catch-rate indices can inform on trends, they do not necessarily help generate understanding of the life history patterns that underpin stock status estimation. Pup survival was the only life history parameter to be estimated in the model, and other parameters such as natural mortality-at-age and the prior mode for pup survival had to be adjusted so that the steepness parameter remained within a reasonable range for the species."

We have covered some of these comments above. The CIE comment about the failure to use age data in the model deserves additional comment in that the model in its present form cannot incorporated size data except indirectly[2], and then it got it wrong in the case of the selectivity curve. There are other stock assessment models

---

[2] Size has to be converted to age, which was done using a von Bertalanffy equation.

available that are able to use size data directly, and it would be useful to employ one or more along with the corrections to the indices and compare results.

The CIE also notes: "Ultimately, the methods used for estimating stock status were found to have been much more sensitive to assumptions about life history parameters than the catch and catch-rate data used in the model." This is a very serious defect. The use of biological parameters in the modeling would be quite useful if these parameters were estimated from data. In practice, only the average number of pups and age at maturity were based on sample data. The former comes from several studies and is consistent with the BLOP data. The age at maturity study, however, is seriously flawed. The animals were not aged, rather length was estimated using a von Bertalanffy equation that may not be correct, and in any event introduces a second source of error that was not accounted for. There is evidence that age at maturity has decreased in recent years. Unfortunately, the study material was discarded, and there is no way to redo the work except collect new specimens.

The other biological parameters used by the model are natural mortality M, pup survival and steepness. Pup survival is estimated by the model, which is a circular process, or fixed by the modeler, which in subjective. The values for M that were decided by SEDAR 11 BW were changed for the final assessment. Steepness was likewise manipulated in order to achieve a credible model output. The fact that these parameters were derived subjectively is disturbing as these are the assumptions the CIE point to as being more influential on the estimate of the status of the stock than are catch and catch-rate data.

3. **Projections**

We did not explore the projections. The future status of the stock is dependent upon the biological parameters, particularly the maturity ogive. Further work needs to be done to include the additional size/age at maturity information and to resolve the inconsistencies in the model results before projections may make sense. In particular, the biological parameters have to be carefully re-examined as they alone control the modeling for the future condition of the stock. Projections need to consider density dependent effects on age (size) at maturity, fecundity and natural mortality. Projections that do not recognize the variability of environmental conditions on growth rate and species interactions such as predation on pups will be misleading over the long term.

4. Conclusions and recommendations

To sum up our conclusions:

- ***The assessment proceeded without using the largest data set available, the BLOP data, which inter alia shows that average age of the catch has not declined over time, as it should if the stock were being overfished.***
- ***The BLOP data also show that the selectivity curve used for the commercial catch is wrong and needs to be re-examined.***
- ***Catch-rates for recent years remain level indicating a population in equilibrium; overfishing is not occurring, whereas the model trajectory indicates a continuing decline in abundance.***
- ***The assessment used several catch-rate series (LPS and NMFS – NE) that were either inappropriate, or did not include the available (but withheld) size and sex data (VA LL).***
- ***The age-at-maturity ogive was derived from a study that is technically flawed.***
- ***The biological parameters used in the model were selected subjectively and there may be some evidence that different values are more appropriate.***

The problem now is that NMFS has used this technically flawed assessment to make the formal finding that the stock is overfished and overfishing is occurring. This starts a legal process that may require a severe reduction in TAC. There is time yet to revisit the assessment before that reduction is in place if NMFS is willing to devote the effort and address most of the concerns the CIE and we have raised. Redoing the maturity ogive study may not fit into this period, but the other work could be done a matter of months.

UNITED STATES DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
NATIONAL MARINE FISHERIES SERVICE
1315 East-West Highway
Silver Spring, Maryland 20910

THE DIRECTOR

Mr. Russell Hudson
Directed Shark Fisheries, Inc.
P.O. Box 11604
Daytona Beach, Florida 32120-1604

JUL 13 2007

Dear Mr. Hudson:

Thank you for your report entitled "Report to Directed Shark Fisheries, Inc. on the 2006 SEDAR 11 Assessment for Sandbar Shark." This document was received by NOAA's National Marine Fisheries Service (NMFS) during the scoping period for an Amendment to the Consolidated Highly Migratory Species (HMS) Fishery Management Plan (FMP). This amendment will implement shark management measures consistent with recent stock assessments.

Based on the February 12, 2003, request of the industry, environmentalists, and academics to improve shark stock assessments, NMFS conducted the 2005/2006 Large Coastal Sharks stock assessment under the Southeast Data, Assessment, and Review (SEDAR) process. This is the same process used by the South Atlantic, Gulf of Mexico, and Caribbean Fishery Management Councils. At all stages, industry, environmentalists, and other interested parties were invited to participate in the process. NMFS believes that the latest stock assessment for sandbar sharks constitutes the best available science.

During the review workshop held June 5-9, 2006, the panel selected by the Center for Independent Experts found that the data and the models employed during the data and assessment workshops, respectively, were the best currently available for evaluating the stock status of sandbar sharks. Your specific concerns regarding which data sets were used in the assessment, selectivity curves employed, appropriateness of catch series included, the age-at-maturity ogive for sandbar sharks, and the selection of biological parameters are addressed in greater detail in the enclosure.

The stock assessment continues to represent the best available science as indicated by independent stock assessment specialists. NMFS is developing proposed management measures consistent with the results of the assessment. There will be additional opportunity for public comments when the proposed rule and draft Environmental Impact Statement for Amendment 2 to the Consolidated HMS FMP are released.

I appreciate your continued involvement in shark management.

Sincerely,

William T. Hogarth, Ph.D.

Enclosure

cc: Dr. Frank Hester
Dr. Mark Maunder

THE ASSISTANT ADMINISTRATOR
FOR FISHERIES

SCIENTIFIC REVIEW OF HESTER AND MAUNDER'S "REPORT TO DIRECTED SHARK FISHERIES, INC. ON THE SEDAR 11 ASSESSMENT FOR SANDBAR SHARK" (2007)

In this section, the Agency addresses each of the bulleted points listed in the report by Hester and Maunder entitled "Report to Directed Shark Fisheries, Inc. on the 2006 SEDAR 11 Assessment for Sandbar Shark" (2007).

• ***1. The assessment proceeded without using the largest data set available, the BLLOP data, which inter alia shows that average age of the catch has not declined over time, as it should if the stock were being overfished.***

This is incorrect. The Bottom Longline Observer Program (BLLOP) data was used in several places during the assessment:

1) Annual average weights from the BLLOP were used to transform commercial landed weights into numbers;
2) An index of relative abundance for the period 1994-2004 was produced; and,
3) Along with other datasets, the BLLOP was used to develop the commercial gear selectivity.

Note that only lengths, not ages, are available from the BLLOP. In the Hester and Maunder (2007) report, length is converted to age using a growth curve from Sminkey and Musick (1995). A lack of declining trend in size through time may be attributed to the fact that the directed bottom-longline shark fishery targets large individuals. While the BLLOP samples the directed bottom-longline shark fishery, one must keep in mind several caveats: the program covers approximately five percent of fishing trips, it was voluntary until 2002, and it is not a fishery-independent index. Fishery independent data is more effective at sampling a larger range of size classes and, therefore discerning trends in size, because any animal that is available to the gear has an equal likelihood of being sampled. Fishery dependent data represents fishermen targeting larger specimens for their greater fin value. Furthermore, fishery-independent data samples random or semi-random locations, whereas, fishermen generally only deploy gear where they have a higher likelihood of catching sharks.

The statement that the average age of a population is expected to decrease under exploitation is true only with qualifiers. For instance, if all ages are fully selected, then exploitation removes the same fraction from each age class over time, so that they are still in relatively the same proportion to each other. In the 2006 assessment, the average age in 1975 is 10.91 yr, and it drops to 9.63 yr in 2004. Over the years presented in their size analysis (1994-2005), the assessment results show that the average age decreased from 9.92 to 9.63 yr. As the age change is not appreciable, it is not surprising that Hester and Maunder (2007) do not find a decreasing trend in their analysis of length trends over years.

Hester and Maunder (2007) further state that the reviewers stopped short of "comparing the model results with actual information from the fishery." Fishery information is included in the assessment in a variety of forms (observer program information, landings, catch), and as such, the model results include fishery information.

Hester and Maunder (2007) indicate that size and sex composition from the BLLOP was not made available at SEDAR 11. This is not entirely correct. The size composition (length-frequency distributions by year and area as well as average length and weight by year) of the catch in the BLLOP was presented in document LCS05/06-DW-16 (Cortes and Neer 2005). Sex information was not presented as the models are not sex-specific. The assessment model assumes males and females are selected by the fishery with equal probability. If there were a sex ratio difference in targeting (biased towards females), this would lead to a greater depletion of mature females than is currently estimated by the model.

Figure 1 in the Hester and Maunder (2007) report, is summarized over all areas, and presumably is intended to support the fact that there was no change in the size composition of the population over the period spanning 1994-2005. However, it is clear from Figures 5 and 6 in the Hester and Maunder report that the size distribution and catch distribution vary by geographic area. Specifically, almost twice as many samples are available from the Mid Atlantic Bight than from the Eastern Gulf of Mexico or the South Atlantic, and the median size and range of sizes is much smaller in the Mid Atlantic Bight. This seems to indicate that the area with the most fishing (Mid Atlantic Bight) has a smaller size composition than areas with less fishing (Eastern Gulf of Mexico, South Atlantic), and likely reflects the impact of fishing on the population.

- ***2. The BLLOP data also show that the selectivity curve used for the commercial catch is wrong and needs to be re-examined.***

As mentioned above, the authors used a growth curve to convert the BLLOP length data to age, and then generated a "catch-at-age" curve for those data. While the BLLOP curve developed differs in shape from what was used in the assessment, the selectivity curve used for the BLLOP was termed "commercial+unreported" catch in the Hester and Maunder (2007) report. It was applied to various gears, not only bottom longline, and was used for a variety of indices (LPS, BLLOP, BLL-logbook, VA LL, PC Gillnet, NMFS LL NE, NMFS LL SE, and Pelagic Logbook). The Catch Working Group at the Data Workshop (in which industry members were represented) decided to reflect the fact that younger age classes can also be retained by (and are available to) various gears. No concerns were raised at that time.

It is interesting to note that, according to Figure 4b in the Hester and Maunder (2007) report, the age of full selectivity developed is actually 5-7 years *earlier* than that of the selectivity used in the 2006 assessment (Figure 4a in the Hester and Maunder report), indicating that more, smaller fish would be available and fully selected by the gear than the current model suggests.

- ***3. Catch-rates for recent years remain level indicating a population in equilibrium; overfishing is not occurring, whereas the model trajectory indicates a continuing decline in abundance.***

The Hester and Maunder (2007) report states that commercial landings and catch rates have remained stable for over a decade, and suggest that this is evidence that the stock assessment results do not reflect reality. It is important to remember that catch rate time series in isolation do not tell the whole story. Most catch rate series show stable or unclear trends in recent years (since the mid-1990s), but the large declines occurred in the late 1970s and 1980s (see VA LL, LPS, and MRFSS). In fact, there has been a commercial quota imposed on the fishery since 1993 (for over a decade); stable landings in the last decade most likely reflect the effect of a commercial quota, not of a stable population. Furthermore, commercial catch declined from 162,000 individuals in 1989 to 72,600 individuals in 1993, prior to implementation of the commercial quota.

Hester and Maunder (2007) also state that the 45% decline in spawning biomass "should be reflected by a decrease in the average age of spawners." However, as clearly pointed out in the Assessment Workshop report (p.86, text following eq. 2), spawning biomass is actually pup production, i.e. $SSB = \sum_a N_a mat_a pup_a$, where mat_a is maturity at age and pup_a is pups produced at age. Thus, the 45% decline in spawning biomass reflects a 45% decline in pups produced, not in weight of spawners. Since all mature sandbar sharks are assumed to produce the same number of pups, the age of the shark is irrelevant in this analysis other than determining whether or not an individual is mature. The 45% decline therefore reflects a decline in abundance (numbers) of mature sharks. The corresponding decline in total abundance for the same period (all ages considered) is 32% (N_{2005}/N_{1994}).

Hester and Maunder (2007) suggest that more extensive sensitivity runs should be made for the recreational catches; however, this point is irrelevant as a catch-free model was also applied to the sandbar shark data (see page 3 in SEDAR11-AW-03) and the same conclusions about stock status were reached.

There is bold text on page 11 of the Hester and Maunder report that discusses the inconsistency between the model predicted decline in SSB (spawning biomass), the lack of decline in average age, and the stable or increasing trend in abundance indices from 1994. As previously pointed out, the model results do not predict a substantial decline in average age, and therefore there is no inconsistency with Hester and Maunder's estimates from the BLLOP data. Eliminating the longer time series disregards critical historical information to help inform the model over the entire history of the fishery. In many cases, trends in relative abundance of sandbar and other sharks have remained relatively stable since the introduction of management in 1993.

- ***4. The assessment used several catch-rate series (LPS and NMFS – NE) that were either inappropriate, or did not include the available (but withheld) size and sex data (VA LL).***

Large Pelagics Survey (LPS)

The Agency does not agree that the inclusion of the LPS data was inappropriate. All decisions to keep or eliminate catch rate series were discussed in detail by the Indices Working Group (in which Dr. Hester participated) during the stock assessment data workshop and a consensus decision was reached. The arguments that the authors make regarding the LPS series (that it has been split in the past due to regulation changes, and that regulations would affect catch) are incorrect. The index was not split in 2002 and the index includes estimates of discards, so the imposition of size and bag limits is not expected to impact the assessment, as there is no indication that fishers would make significant changes to targeting methodology. Finally, contrary to what is stated: "The LPS index is clearly two essentially flat indexes" is not true, as the 1985-1992 period shows a declining trend.

The text on page 9 of the Hester and Maunder report (2007) referring to the LPS index is inconsistent with earlier statements on page 7. Specifically, on page 7, Hester and Maunder (2007) summarize the model trajectory: "The slow decline in SSB/B0 [sic] between 1975 and 1981 shown in Figure 7 results from the recreational catches, which are the only appreciable catches assumed. Catch is the only thing that makes this model decline as there is no annual random variation in recruitment (and no catch-at-age data to estimate it). Recreational catch is believed to target young sharks and therefore some time must elapse before the effect of taking young fish shows up in the biomass of older fish. The commercial fishery, which targets larger fish, begins in the mid 1980's and, combined with the effect of the removals of the younger fish earlier on, is followed by an immediate and more rapid decline in SSB/B0 [sic] reaching a depletion level of 0.31 in 2004." This is a reasonable summary of the model behavior, yet on page 9, Hester and Maunder (2007) write that "The LPS index (Figure 10 in the Hester and Maunder report) has the same difficulty as the VA LL index in that it consists of two periods with high values in the early period and lower values in the recent period." The Agency questions why this is characterized as a *difficulty* when it is completely consistent with the earlier text describing the catch history and the impact on SSB. Hester and Maunder (2007) state that the index should be split as it had been in the past to reflect the change in regulation (size and bag limit). Two points need to be clarified here: 1) the LPS index was not split in 2002 and 2) the LPS index includes estimates of discards, so the imposition of size and bag limits is not expected to impact the assessment as there is no indication that fishers would make significant changes to targeting methodology.

Virginia Institute of Marine Science Longline Survey (VA LL)

Hester and Maunder (2007) state that the lack of shark age information from the Virginia Institute of Marine Science's longline (VA LL) survey led the Agency to inappropriate size selectivities. While length information from the Virginia Institute of Marine Science's longline (VA LL) survey was not incorporated in the current assessment (only one series containing all sizes was developed), it may not have provided any additional useful catch rate series had it been available. The data would have to have been converted from length to age, and then further separated into the various life stages (neonate, juvenile, adult). Since the entire data set is relatively small, the model may not have been able to standardize the series because of the limited number of samples. Breaking larger data sets into smaller pieces may have resulted in there not being enough information to conduct the analyses required. The various age/stage series used in the 2002 assessment were nominal. It is likely that the assessment participants might not have been able to develop standardized indices by stage because as one partitions a small data set into even smaller parts, one may end up with insufficient observations for all years to conduct the statistical standardization.

Hester and Maunder (2007) also state that the VA LL index omitted data prior to 1981, which is incorrect. The data included years from 1975 to 2004. Furthermore, we note that this index was used without contest in the 2002 assessment. Additionally, Hester and Maunder (2007) state that age data contained within the VA LL data set was not made available. As stated previously, there is no age data for the majority of the time series, just length information; the exception being when directed ageing studies were being undertaken, but these constitute only a very small portion of the total sample.

Hester and Maunder make the statement on page 11 that "The VA LL index may or may not be usable when and if it is properly standardized." The Agency is not sure to what the authors are referring. This index was initially standardized using a general additive model (GAM). Following recommendations of the Indices Working Group, the index was re-standardized for consistency with other indices following the Lo et al. (1992) methodology. Therefore, it was properly standardized, using well established and accepted statistical techniques.

The authors also state that, with regards to the VA LL index, "it is impossible to know what size or sex selectivity to apply to the series" since size or sex data were not provided for review during SEDAR 11. The 2002 assessment used several nominal forms of this index, split by stage/size class. At that time, all sizes/stages were represented, thus a selectivity representing all stages is appropriate for use in the current assessment as all stages/ages would still be represented in the data. Additionally, sex-specific selectivities are not utilized for any index.

National Marine Fisheries Service Northeast Index (NMFS NE)

The text on page 10, in particular in the legend for Figure 11, suggests that the high uncertainty for the NMFS NE index, and the fact that it has only 4 points should have precluded it from being used in the assessment. Giving all indices equal weight or weighting all indices by their CVs resulted in the same conclusion. The statement by Hester and Maunder (2007) that the NMFS NE index "has no statistical trend other than zero, but when given equal weight in the assessment has considerable influence on the outcome" is false. It contributes very little to the objective function (Assessment Workshop Report. page 122, Fig. 4.12).

- ***5. The age-at-maturity ogive was derived from a study that is technically flawed.***

The Agency does not believe this to be the case. Hester and Maunder (2007) state that the study is "technically flawed" since the animals used in the reproductive study were not directly aged, but rather length was converted to age using a von Bertalanffy growth equation. The Agency would like to make several points in this regard:

1) This is commonly the case with reproductive studies, especially when an age and growth study already exists for the species;
2) The impetus for using an existing age and growth study for conversion from length to age is especially strong in this case because the specimens used in the age and growth study (Sminkey and Musick 1995) were collected within approximately 5 years of the reproductive samples, and were collected from the same region; and,

3) Hester and Maunder (2007) used this very same growth curve (Sminkey and Musick 1995) to convert the length information from the BLLOP to age for use in deriving a selectivity curve for those data.

Hester and Maunder (2007) criticize the maturity study as being flawed. "The animals were not aged, rather length was estimated using a von Bertalanffy equation that may not be correct, and in any event introduces a second source of error that was not accounted for" (page 15 of the report). Despite making this comment, Hester and Maunder (2007) begin this subsection by saying "Converting lengths to ages using a von Bertalanffy equation (Sminkey and Musick, 1995) gives the distribution for the BLLOP catches shown in Figure 3" (page 4 of their report). It would seem that Hester and Maunder (2007) are arguing both *for* and *against* using the von Bertalanffy growth equation which is logically inconsistent. By using the same von Bertalanffy equation that Hester and Maunder (2007) criticize, they are also introducing the same "sources of error" that they advise against introducing.

Furthermore, the Sminkey and Musick (1995) growth curve was constructed based on animals caught off the coast of Virginia. Specimens for the maturity ogive (Merson 1998) were also collected from this area within five years of when the specimens for Sminkey and Musick (1995) were collected. Hester and Maunder (2007) state that the BLLOP dataset is "representative of about 90% of the commercial catch of sandbar" but "does not include the North Atlantic Region (which includes Virginia)" (page 2 of their report). Given the difference in geographic coverage between the BLLOP samples compared to the Sminkey and Musick (1995) growth curve they use to convert lengths to age, it would seem that the Merson (1998) and Sminkey and Musick (1995) studies employed in the stock assessment would be more compatible since they both employ animals caught from the same region.

- ***6. The biological parameters used in the model were selected subjectively and there may be some evidence that different values are more appropriate.***

This is incorrect. The values of life history parameters agreed to at the data workshop produced steepness values below the theoretical minimum level. A document produced at the assessment workshop, SEDAR11-AW-10 (Brooks and Cortes, 2006), reflected the range of life history parameters that could be adjusted to raise steepness, and the biological argument that could be made for each change. After much discussion at the Assessment Workshop, the group decided that adjusting the values of M (natural mortality) was the most justifiable action. The Agency has emphasized that steepness was not "manipulated to achieve a credible model output" as Hester and Maunder state. Rather, steepness is a function of all life history parameters (natural mortality, maturity, and fecundity—see SEDAR11-AW-10). None of these parameters were derived subjectively; rather, they were estimated based on ecological principles consistently applied to a variety of species. We are unaware of the "evidence that different parameter values are more appropriate." If this "evidence" exists, it was not provided during SEDAR 11 for discussion. The same logic for modifying the biological parameters was applied to blacktip sharks; however, in that case, the logic was deemed acceptable by Hester and Maunder (2007).

In the 2002 assessment, a similar problem was encountered with steepness estimates falling below the minimum bound. As indicated in the 2002 assessment report, the parameter for pup production was inflated, yet this approach was not contested during the 2002 assessment. Adjusting pup production was presented as an option this time, but it was rejected by the group in favor of adjusting natural mortality levels.

Hester and Maunder (2007) also mention in the Reviewers comment that "No age data were used in the model." This was due to the fact that catch-at-age data are not available for sandbar sharks. One cannot simply convert length to age based on an age and growth study for use in the model, but must derive catch-at-age data from ageing samples collected from the fishery. This type of data does not currently exist for this or any other species of shark exploited in U.S. waters.

While the authors state that "there is evidence that age at maturity has decreased in recent years", this "evidence" was not presented during SEDAR 11. A chapter in a Ph.D. dissertation by Rebeka Rand Merson (1998) investigates reproduction of sandbar sharks and was reviewed by the Life History Working Group at the Data Workshop and determined to be the best available data. This chapter was included in SEDAR 11-DW-47. Concerns stated by Dr. Hester during the SEDAR data workshop were addressed by Merson in SEDAR 11-AW-09 during the assessment workshop. The Merson study is the only study available which provides a maturity ogive for sandbar sharks in the region of interest.

Hester and Maunder (2007) had further issues with the biological parameters used in the model. They state that "Pup survival is estimated by the model, which is a circular process." The Agency does not believe this to be true. While pup survival is given a prior probability distribution, it is estimated by the model - there is no "circular" process involved.

RESPONSES TO OTHER CONCERNS RAISED IN THE REPORT

This section provides additional detail and discussion of other issues presented in the Hester and Maunder (2007) report.

Indices Employed in the Stock Assessment

The issue of consistency of indices over time was discussed to some extent by the Indices Working Group at the Assessment Workshop. Standardization is used to account for changes in factors, such as introduction of regulations. While regulation changes could affect an index, this is not the case for the VA LL index, which is fishery-independent, and the LPS, where total catch (*i.e.,* not just landings) is recorded. The change from voluntary to mandatory observer coverage in the BLLOP is more of an issue and could have affected the standardization of that index more than the two indices mentioned by Hester and Maunder.

Furthermore, Hester and Maunder (2007) refer to "good" and "bad" indices that were employed, which is a very subjective characterization. Fisheries-independent indices typically have small sample size, which often results in a higher CV than fisheries-dependent indices that are derived from thousands of observations. The base case treated all indices as having equal weight, but a sensitivity model using the CVs for each index was also run, and arrived at the same conclusion that the stock is overfished with overfishing occurring. Hester and Maunder correctly point out that all factors in a standardized index are assumed to have been constant over time, and that if this is not true, then the resulting index could be biased. This point is equally true for fisheries-dependent indices for instances where catchability has changed over time either through increased skill at finding fish, learning, or gear and technology improvements, etc.

Hester and Maunder (2007) present "trend analyses" for several of the input indices. These fits conducted by the Hester and Maunder report appear to have been done in the EXCEL software. However, their analysis is misleading and misapplied. First, the indices fit within the assessment model assume a lognormal error structure, which was assumed to be the most appropriate error structure to use, while EXCEL fits trendlines assuming normally distributed errors, which could contribute to the differences seen in Figure 1 in this document. A comparison of the model fits versus "trendlines" to the three indices that Hester and Maunder single out (VA LL, LPS, and NMFS NE) is shown in Figure 1. For the VA LL and LPS indices, the trendlines of Hester and Maunder are far more negative than those predicted in the assessment model; the NMFS NE assessment predicted index is fairly similar to the trendline fit. The model fits to the observed index values are shown in the assessment final report, pages 112 - 114, Figure 4.7. From those plots, and from the likelihood contributions of each index (Assessment Workshop report p.122, Fig. 4.12), it is clear that the LPS is poorly fit and that it is not driving model results.

Sensitivity Analyses

From pages 11 to 15 in the Hester and Maunder report, various sensitivity model runs are discussed. Despite earlier comments by Hester and Maunder about "bad indices" and "negative slopes," excluding or down weighting those same indices did not alter model results substantially. In Table 1 of the Hester and Maunder (2007) report, they summarize model results for SSB_{2004}/SSB_0 rather than SSB_{2004}/SSB_{MSY}. Fishery management in the United States has determined the overfished criterion based on MSY reference points and not on relative depletion (we note that they correctly summarize the overfishing statistic as F_{2004}/F_{MSY}). This misunderstanding of reference point seems to be carried even further, as Hester and Maunder appear to believe that 50% of SSB_0 is the benchmark determining whether a stock is overfished or not. This is not the case. The correct benchmark is SSB_{MSY}.

Run 2 (No VA LL): Again, the years prior to 1981 in the VA LL index were not omitted, and the indices used in 2002 were nominal only. In all these sensitivity analyses, it is stated that the index of interest was "heavily down weighted", but there is no explanation of how exactly it was done, other than saying that "this is about equivalent to using variance weighting" (stated under Run 3 (No LPS)). The naming convention used is confusing: "No VA LL index" or "No LPS" implies the index was removed from the run, but that is not the case according to the text provided.

Run 5 (Removing VA LL and LPS): By selectively removing the 2 indices that start earlier in time (1975 and 1986, respectively), the model essentially has no CPUE information prior to 1993, thus showing a much reduced level of depletion since there are no points of comparison prior to 1993.

Run 7 appears to be in agreement with the base case assessment results, although the authors suggest that further sensitivity runs should be made to explore the poorly estimated recreational catches. Again, we note that the catch free model (which does not use any catch) arrived at the same results. The catch free model is discussed in SEDAR11-AW-03 (Brooks 2006).

The Hester and Maunder sensitivity analyses, in most cases, support the original stock status determined by the assessment. The main points that the industry had taken issue with (*i.e.*, the VIMS LL series, the LPS series, the commercial selectivity, and the median age at maturity) made essentially no difference, as demonstrated by Hester and Maunder's (2007) own sensitivity runs. The only difference was obtained when completely removing or down-weighting the 2 long-term CPUE indices (or other extreme scenarios), thus removing critical historical information from the modeling. Also note that for all scenarios (except run #12 where pup survival = 0.85) in which no overfishing ($F_{2004}/F_{MSY}<1$) was found, pup survival was unrealistically high (0.93-0.97).

CONCLUSIONS

While raising important concerns, the Hester and Maunder (2007) report contains numerous factual errors and unsupported statements. For the reasons stated below, none of the 6 initial conclusions (bullets in the abstract) stated in the Hester and Maunder (2007) report are valid:

1) The BLLOP index was used at different steps in the assessment process;
2) The new selectivity created by Hester and Maunder using only BLLOP data ultimately did not have any effect on results;
3) There has been a commercial quota imposed on the fishery since 1993 (for over a decade) and stable landings in recent years most likely reflect the effect of quotas, not of a stable population;
4) The CPUE indices used in the model were appropriate and selected by a group of experts and endorsed by a peer-review panel. Selective use of CPUE series will eventually produce the desired results, however, this could bias the outcome;
5) There is no evidence to suggest that the maturity/reproductive study is technically flawed; on the contrary, it is the best available science, to date; and,
6) The biological parameters used in the model were sensible and biologically defensible, unlike those in the optimistic outcomes produced by the Hester and Maunder (2007) report. There is no evidence to suggest that other more appropriate values are available or should have been used in the assessment.

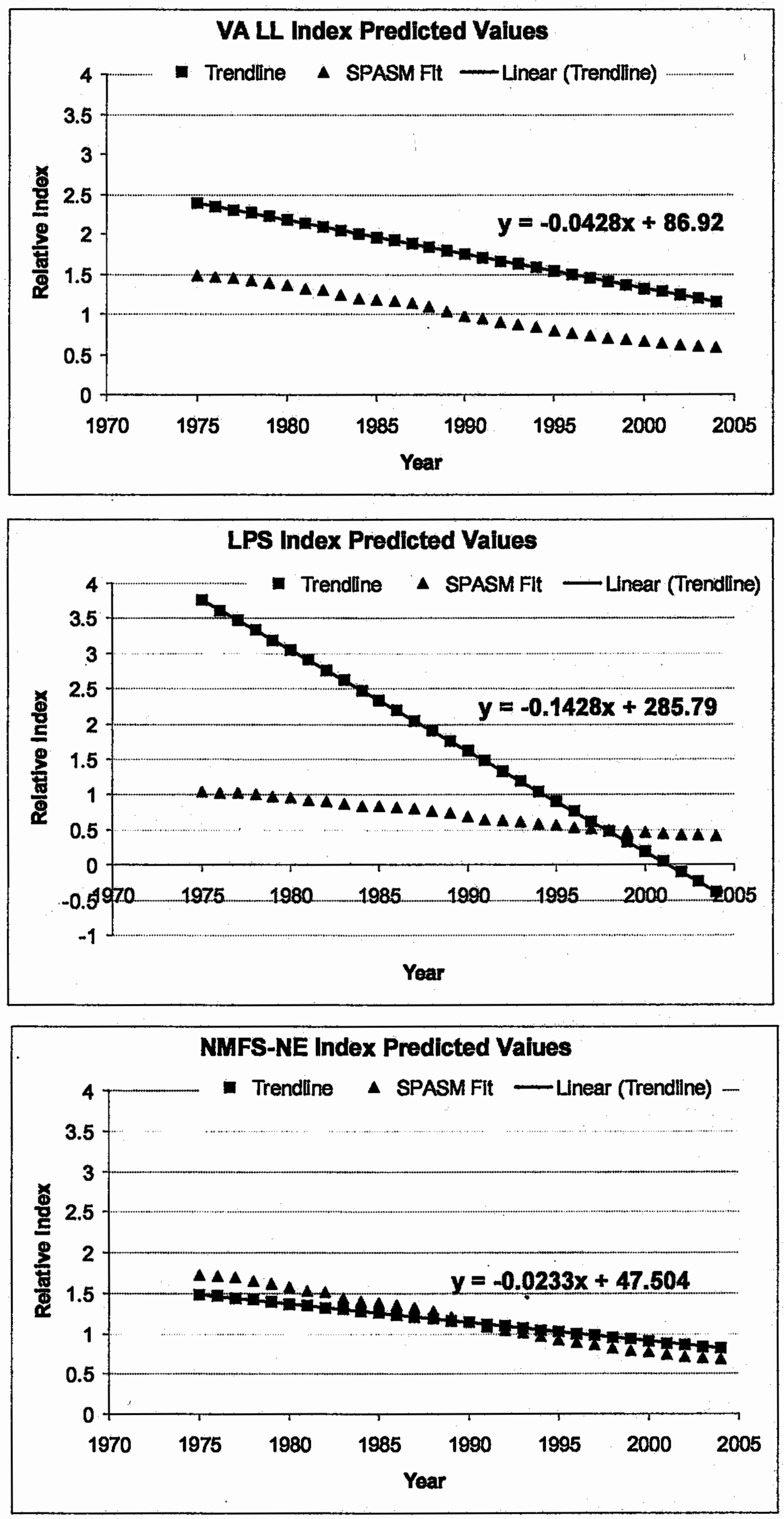


Figure 1. Comparison of model-predicted indices (solid triangles) to the trend lines of Hester and Maunder (solid line with solid square symbols).