Report on the 2004 Assessments of Black Seabass, Sea Scallop and Atlantic Bluefish in the Northeast United States

Michael J. Armstrong

Prepared for

University of Miami

Independent System for Peer review

The Centre for Fisheries and Aquaculture Science
Lowestoft Laboratory
Pakefield Road
Lowestoft
Suffolk NR33 0HT
England, United Kingdom
Phone +44 1502 524362
www.cefas.co.uk

CEFAS
The Centre for Environment, Fisheries & Aquaculture Science

Contents

Executive Summary

The 2004 assessments of Black Seabass, Sea Scallop and Atlantic Bluefish in Northeast United States waters were reviewed as part of the SARC 39 (Stock Assessment Review Committee No. 39) process. The Assessment Review Panel met at Woods Hole, Massachusetts from 7-10 June 2004. The assessments of the stocks were presented to the Panel and the validity of the data, assessment procedures and results were discussed. A proposed new assessment method for Sea Scallops was presented and discussed. The Panel Members then prepared their individual reviews.

The Black Seabass data and assessment are considered adequate for evaluating stock status relative to agreed reference points. A well-designed and executed tagging programme has indicated that fishing mortality is likely to be below $F_{MAX,}$ whilst survey data indicate a growth in biomass since the late 1990s, although indices are declining again.

Trends in Sea Scallop abundance are well estimated from intensive dredge surveys and appear robust. Overall, abundance is above the targets and thresholds specified in the Fishery Management Plan although there are regional differences in recent trends with Georges Bank stocks now declining and Mid Atlantic Bight stocks continuing to increase due to strong recruitment. Fishing mortality is less well estimated, and whilst it appears to be above the F_{MAX} threshold when averaged over all regions, the estimates appear to be relatively low on the Georges Bank. The latter conclusion may be sensitive to trends in cull size and discarding.

The Atlantic Bluefish assessment is not considered adequate for use by fishery managers. A mistake was found in the calculation of trends in catch per unit effort, and an inappropriate survey series was used in the production model.

Some key recommendations are summarised below:

- Investigate methods for extracting more accurate recruitment signals from multiple surveys of seabass and bluefish
- Collect adequate age composition data for surveys and commercial catches of black seabass and bluefish to allow age-based assessment.
- Obtain better estimates of discarding, including size/age composition, survival rate of discards, and quantities discarded.
- Further develop tagging studies to investigate mortality, selectivity and migration of black seabass and bluefish
- Further develop the CASA model for scallops, as well as more parsimonious methods applied at smaller spatial scales
- The use of recreational CPUE as an index of population size needs to be thoroughly evaluated.

1. BACKGROUND

This report reviews the 2004 assessments of black seabass, bluefish and sea scallops in northeast US coastal waters, at the request of the University of Miami (see Appendix 5). The author was provided with draft stock assessment reports and web access to relevant files and documents (Appendix 4), and participated in the 39th Northeast regional Stock Assessment Workshop (SAW39) Stock Assessment Review Committee (SARC39) Meeting.

2. REVIEW ACTIVITIES

The SARC 39 meeting was held at the Aquarium Conference Room - Northeast Fisheries Science Center, Woods Hole, Massachusetts from 7-10 June 2004. The Panel membership is listed in Appendix 1. The agenda for the meeting is in Appendix 3.

The meeting was the first of the “new model” of SARC reviews with a smaller panel than previously, although with the same number of invited reviewers. The meeting was open, and was attended by observers including members of the fishing industry. The draft assessment of each stock was presented to the Panel and other attendees, and the input data, models, parameter estimates and biological reference points were evaluated through open discussion. A conclusion was then drawn on whether to accept the assessment as a basis for management of the fisheries. The Terms of Reference for each stock (Appendix 2) were reviewed to ensure they had been fully addressed, and recommendations from the previous SARC report were reviewed to determine the extent to which they had been addressed.

3. FINDINGS

3.1 Black Seabass

3.1.1 Summary

The assessment presented at SARC39 is considered of acceptable quality as a basis for fishery management for the following reasons:

- The Spring NEFSC survey index of biomass for fish greater than or equal to 22cm long remains well above the threshold of 0.5 * mean index for 1974 –1976 (taking the 3-year mean for these years as a proxy for B_{MSY}), and there has been no change to survey design or protocol since previous assessments.
- Considerable research effort has been devoted to carrying out a well-designed and implemented tagging programme, and the results strongly indicate that recent rates of fishing mortality are below the currently adopted F_{MSY} proxy (F_{max} =0.33 from yield per recruit).

Additional analyses using the Index Method from the NOAA toolbox indicate that recent fishing mortality is less than required for the stock to replace itself, and hence that further growth of biomass is possible.

The population occurs along the outer (deep water) boundary of the survey during spring, and hence could be subject to changes in availability to the survey caused by environmental influences on cross-shelf distribution. Whilst this could explain the periodic increases in the survey index, a progressive shift in the composition of landings towards “large” and “jumbo” categories since the late 1990s suggests the influence of strong recruitment rather than changes in availability to the survey.

3.1.2 Comparison with previous assessment (SAW 27)

The previous assessment reviewed in 1998 presented the time series of survey indices and also used simple analyses of commercial and recreational length frequencies to estimate fishing mortality. These estimates were compared with F_{MAX} from yield-per-recruit. The present assessment (SAW39) extends the survey time series, but focuses on the results of a new tagging study to estimate mortality rates.

3.1.3 Fishery data used for the assessment

Fishery data are only used for application of the Index Method, which calculates a proxy for fishing mortality as the ratio of annual fishery catch to a 4-year running mean of the survey index.

The accuracy of commercial landings data has improved over time due to changes in the reporting system. However, no information was presented on discarding in the commercial fishery. Recreational catch has been at a similar level to commercial landings, but is far more variable from year to year. It was suggested that this might reflect inaccuracies inherent in the survey method for estimating recreational landings and discards. The large variations in the 1980s, with very high recreational landings in 1982 and 1986, were a particular cause for concern. The precision of the catch estimates should be presented. In discussion, the coefficient of variation of the recreational estimates was cited as around 20-30%.

It was noted that escape vents in pots limit discards. Larger discard rates occur in the trawl fisheries, particularly small mesh fisheries for squid.

3.1.4 Fishery independent data used for the assessment

Numerous research survey series were presented, some limited in spatial extent, others covering a larger part of the northeast US coast. Autumn surveys were considered less reliable due to associations between black seabass and underwater structures in inshore waters at this time. Winter-Spring surveys consistently show an increase in abundance since the late 1990s, although with different years of peak abundance. In most cases the recent peak has been followed by a sharp decline.

Despite the multiplicity of surveys, only one (NEFSC Spring) was used in the assessment because it covers a large area with consistent methods over time. Attempts should be made to extract as much information as possible from all the series considered appropriate, using, for example, a GLM or GAM approach to combine the various surveys into a standardized index. This is a broader issue covering species such as bluefish as well, and could valuably be addressed through a Workshop attended by State and Federal scientists.

A potential area of concern is the cross-shelf migration of black seabass resulting in the spring survey catch-rates being highest at the outer boundary of the survey grid. If the stock distribution extended beyond the survey to different extents in different years, or if the vertical distribution changed along the shelf edge in response to environmental conditions, strong year-effects would be apparent in the survey. Without age-composition data, it is difficult to evaluate if the periodic occurrence of large catch-rates (as apparent since the late 1990s) is an effect of survey catchability or strong recruitment.

Survey indices for juveniles show poorly correlated fluctuations that are not consistently associated with the changes in Spring survey biomass indices, and the mean length of black seabass in the Spring NEFSC offshore survey varies considerably from year to year without any trend in recent years that might suggest an expansion of age structure due to a series of strong recruitments. Commercial fishery data provided late in the SARC39 review meeting showed a progressive increase in the percentage of landings categorised as “large” or “jumbo”, suggesting an increase in average size of fish in the stock that is not reflected in the survey. This could also reflect high grading.

It is recommended that the survey indices be disaggregated by age to identify the impact of year class variation in the biomass index and to investigate the magnitude of year effects. The relationship between offshore distribution patterns, environmental indicators such as temperature fronts, and timing of surveys, should also be investigated, as part of the recommended workshop on surveys.

Although the biological reference point for biomass is based on survey data derived from non-transformed catch-rates, many of the series in the SAW39 report give geometric mean catch rates. This appears to have been done to avoid problems with some stations having very high catch rates that dominate the overall index. It was noted that the mean indices were bias corrected after back-transformation. Transformations are not commonly applied in eastern Atlantic surveys to derive survey indices for use in assessments. The theoretical justification for applying transformations should be reviewed.

3.1.5 Tagging programme

The scientists involved in the tagging study are commended for initiating a very valuable and well-structured program in response to the previous SARC recommendations, and are encouraged to continue with this approach to establish variability in mortality estimates.

The calculations of mortality rates involve several variables measured with error. These include reporting rate, rate of tag loss and proportion recaptured. It would be useful to derive the variance of the estimated exploitation rate based on the likely variances of these parameters.

The assessment report rightly points out that an assumption of tagging models is that a tagged fish should have an equal probability of being recaptured as a similarly sized non-tagged fish in the population, in order to calculate unbiased estimates of fishing mortality. However the report does not really address the possible biases in estimates due to not meeting this criterion.

The fifth figure in Appendix 1 of the SAW39 report shows a cluster of tag returns within 50 days of tagging in Fall 2002, during which time many fish had not dispersed very far. Any fish recaptured within 7 days were ignored in the calculations of exploitation rate, but the reason for choosing 7 days to censor the data is not explained, the numbers of recaptures in this period are not documented, and the implications are not discussed. To avoid bias in the Petersen model estimates, it would be necessary to reduce the numbers tagged by the numbers recaptured within 7 days, adjusted for tagging mortality. It is not clear if this was done, although the effect of ignoring this may be small.

I would agree that trying to derive complicated weighting factors to get an overall mean F from regional tagging data is fraught with difficulties, and the attempts by the Working Group to look at sensitivity to this is commended.

It is recommended that tagging continue for a sufficient time to establish the stability of estimates of exploitation rate, and to allow more sophisticated analysis methods.

3.1.6 Results of the assessment

The assessment of current stock biomass relative to historical values is inferred from the Spring NEFSC survey results for black seabass greater or equal to 22cm in length. There is evidence for declining survey indices in winter-spring surveys in the last couple of years, following some high values up to 2002. The dependence of management decisions on raw, age-aggregated survey indices is a serious shortcoming as there is no way to distinguish year-effects in catchability from genuine changes in biomass. Without robust data on year-class variations, it is also difficult to make useful forecasts, particularly if biomass variations are strongly driven by recruitment due to a truncated age composition.

Fishing mortality rates from the 2002 and 2003 tagging experiments were relatively close, and below F_{MAX}. The experiments appeared well designed and executed. However, the accumulation of further similar estimates from repeat experiments would give greater confidence in the results.

3.1.7 Biomass and fishing mortality reference points

The reference point for biomass appears to have no analytical basis, and has been set at the average of three successive biomass indices close to the highest in the series. As a result, the biomass index is below the threshold in most years. The fishing mortality threshold is set at F_{MAX} from a yield-per-recruit analysis. As there is no age-based assessment, the origin and reliability of the selectivity pattern used in the YPR analysis is not clear. The F_{MAX} estimate is not necessarily compatible with the tagging estimates of F. Some of the tag returns for fish above the cut-off of 28cm were probably 3-year-olds or younger, which have selectivity < 1.0 in the YPR analysis (mean weight of 3-year-olds in the catch is given as 0.321 kg, whilst a 28 cm fish is likely to be closer to 0.2 kg).

3.1.8 Recommendations of previous review

The SAW 27 report contains the SARC comments and a list of research recommendations. The latter were: 1) increased at-sea sampling; 2) obtain commercial length frequency data, by market category, from North Carolina from 1984-93 and 1997; 3) initiation of a tagging programme; 4) ageing should be updated to include the most recent samples; 5) further study of size/age and density effects on sex changes, and sex-specific mortality rates and growth; 6) determine the value of artificial reefs for increasing fish production; and 7) consider the utility of pot surveys because of potential catchability problems in trawl surveys for a species that aggregates on hard structures.

The SAW 39 assessment deals mainly with recommendation 3, with updated catch and survey data. Recommendations 5 to 7 are not considered, and indeed the whole issue of sex change and artificial reefs is ignored.

3.1.9 Recommendations for future black seabass assessments

- More comprehensive evaluation of regional survey data is required to give more integrated indices of recruitment. For example, catch rates of recruits could be modelled as a function of location, time of year and gear type in the surveys to provide standardised indices. Good understanding of recent recruitment dynamics is essential for forecasting, as this stock appears to have quite variable recruitment, and periodic population growth may be due to good recruitment.

- Age-structured indices should be derived for the winter-spring offshore surveys to allow better interpretation of the survey results, including identification of year-effects and year-class effects.

- The accuracy and completeness of catch data, particularly recreational catch, should be investigated to explain the unusual interannual variability.

- Confidence limits for survey-based estimates of recreational catch should be presented.

- More information on discarding practices, and factors affecting discarding, is required.

- The tagging studies should valuably continue to allow return rates over longer periods to be established. This would require longer-term data on tag loss rates.

- The relationship between offshore distribution patterns and environmental variables such as temperature and frontal systems should be investigated to ensure that catchability effects are not driving trends in the spring surveys.

3.2 Sea Scallop

3.2.1 Summary

Trends in Sea Scallop abundance are well estimated from intensive dredge surveys and appear robust. Overall, abundance is above the targets and thresholds specified in the Fishery Management Plan although there are regional differences in recent trends, with Georges Bank stocks now declining and mid Atlantic Bight stocks continuing to increase due to strong recruitment. Fishing mortality is less well estimated, and whilst it appears to be above the F_{MAX} threshold when averaged over all regions, the estimates appear to be relatively low on the Georges Bank. The latter conclusion may be sensitive to trends in cull size and discarding. Further development of length-based assessment models (CASA) for scallops is encouraged, as well as investigating more parsimonious methods applied at smaller spatial scales than CASA.

3.2.2 Comparison with previous assessment (SAW 32)

The previous assessment used catch-rates in dredge surveys to monitor biomass, and estimates of fishing mortality from ratios of commercial fishery catch to survey biomass. The latter were re-scaled to the long-term mean of annual estimates of F calculated from catch-rates (by number) of pre-recruit and post-recruit scallops in the same survey series. Estimates of dredge efficiency were derived, and results of a new length-based yield-per-recruit model were presented.

SAW39 updated the survey-based estimates of fishing mortality and biomass trends for the Georges Bank and Mid-Atlantic Bight, and presented some estimates of natural mortality using “clapper” ratios. Estimates of gear selectivity were given. Various ancillary studies were given as appendices to the SAW39 report, including evaluation of errors in SMAST video estimates of scallop shell heights, evaluation of the effect of rock chains on catch-rates of survey dredges, and calculation of dredge efficiency using SMAST video estimates of scallop density. A detailed presentation was given on an implementation of Catch At Size Analysis (CASA; Sullivan *et al*. 1990).

The updated survey biomass estimates for the mid Atlantic Bight show continuation of the high abundance recorded for the final years of the SAW 32 assessment, with a very large increase in both the open and closed areas in 2003. SAW 32 showed increased biomass on the Georges Bank in the late 1990s, particularly in the closed areas. The SAW39 report shows a decline in the open and closed areas of the Georges Bank since 2000.

The SAW 39 report uses the same method for estimating fishing mortality as adopted by SAW 32. However, there are differences in the historical survey estimates, presumably due to re-working of indices. Causes of amendments to any such input data files should be explained, if only to distinguish between true updates and errors in transcribing data. The re-scaled F estimates for the Georges Bank indicate more than an order of magnitude decline in F from the late 1990s compared with the 1980s, whereas in the Mid Atlantic Bight, the estimates of F show only a 2-fold to 3-fold decline over time.

3.2.3 Fishery data used for the assessment

Landings data are used for calculation of fishing mortality indices, whilst length frequency distributions (LFDs) of landings were used in the proposed CASA model. The landings data have a number of sources of error, compounded by poor information on quantities and sizes of scallops discarded over time. Quality of data appears to have improved since the mid 1990s when a mandatory reporting system with dealer reports and vessel trip reports was introduced.

During 1982-1992, when meat count regulations were in force and recruitment was high, there was probably an increase in unreported catch. Further, during the late 80’s and early 90’s there was general category fishery in the Gulf of Maine in which scallop catches may not have been reported. There was probably also an unreported catch in the flounder fishery. The non-reporting of catches will cause under-estimates of fishing mortality using the catch:biomass method.

Fishery discards affect the catch data used both for the current estimates of fishing mortality and for the proposed CASA method, which at present excludes discards data. Observer data show an increase in cull size and proportion discarded from the early 1990s. It is assumed that 20% of the discards die, but this may vary considerably, particularly if scallops are not returned immediately to the water in hot weather or if they have been retrieved through surface layers of water at lethal high temperatures in summer. On the other hand, fishing industry representatives suggested that the figure of 20% discard mortality was too high. They also emphasized that fishermen are now utilizing 4-inch rings to avoid catch of small scallops, which has reduced discarding.

It was reported that there were problems with sampling for LFDs prior to 1994. Agents would measure shells in “the last basket” taken and saved by the boat. However, meat-count regulations affected fisherman’s behaviour. On the way back to port they might stop to fish in areas with large scallops despite low catch rates. These behaviours and regulations can introduce bias to length composition data. Sea-going samplers have measured random samples of catch since 1994. Sea samples were compared to port samples for 1994 and differences were substantial. Therefore, only the sea-sampler LFDs have been used since 1994. However, in some years there are not enough trips sampled. Fishery LFDs for the period 1985-1993 were excluded from the CASA model.

The sources of error in the fishery catch and LFD data will contribute to different forms of measurement error in the assessment procedures. This is an important consideration for the estimation of fishing mortality using the catch:biomass method, which has not been formulated as an observation-error model with explicit treatment of the measurement errors inherent in the data. A comparatively poor fit of the CASA model to some commercial LFDs could indicate the relatively poorer quality of the commercial fishery data compared with the survey data.

3.2.4 Fishery independent data used for the assessment

In comparison with some of the fishery data, the scallop dredge survey data are of very high quality. The surveys appear well executed, and extensive studies of survey gear selectivity and efficiency have been carried out. Information from video surveys carried out by SMAST give further confidence to the results. The ability to accurately monitor both the size composition and catch rates of surveys in closed and open areas provides a powerful assessment and management tool. This potential is somewhat diminished by the Fishery Management Plan requirement to classify the status of the combined northeast stocks irrespective of trends in different areas.

3.2.5 Life history parameters used in the assessment

The basic assessment method requires only an estimate of natural mortality. Growth parameters are required for the length-based yield per recruit and for the proposed CASA method.

A novel method of estimating natural mortality using clapper ratios was presented. Earlier studies reported at the meeting indicated a natural mortality rate of about 0.1 yr^{-1} although data provided by the working group showed lower values interspersed with periodic increases that could be related to effects of changes in size composition, or to episodic mortality events. The assessment scientists were content with the use of M=0.1 in assessments, for scallops > 40mm shell height.

The effect of predation by starfish and lobsters could vary substantially between areas and time periods. This could be important at the smaller spatial scales of closed areas. Studies on predation rates in areas covered by the dredge surveys could provide useful information to help interpret apparent changes in abundance.

There are some indications of faster growth rate in closed areas (possibly 30% faster). Scallops spend their lives where they settle and growth depends on local conditions including depth. Interpreting such differences in growth would require consideration of how fishing mortality and gear selectivity affect length-at-age distributions. Changes in growth rate could also occur due to environmental drivers such as temperature and plankton production. An increase in bottom temperature at the southern part of the scallop's range was reported, and the scallops were reported to be faring more poorly there. In view of the move towards length-based assessments that require growth parameters, better understanding of spatial and temporal variations in both the mean and the variance of growth rates, and the causes of such variation, is needed.

3.2.6 The SAW39 assessment procedure

The monitoring of biomass using scallop dredge surveys is likely to provide robust data on stock trends for use in the Fishery Management Plan. Bootstrapped confidence limits appear quite tight, particularly for the Mid Atlantic Bight. Some means of evaluating year-effects in surveys (e.g. due to undetected change in dredge efficiency), as might be shown by anomalous catchability across all sizes, would be advantageous. The assessment team must consider if applying a relatively complex model such as CASA, which will introduce smoothing through assumptions of constant growth and natural mortality over time, and which also fits to poorer-quality commercial fishery data, is able to provide more accurate stock trends than the raw survey data.

The catch:biomass, 2-bin and re-scaled F methods of estimating fishing mortality have remained essentially unchanged since the previous assessment, although there are some small changes in the input data between assessments. There are a number of problems with the method:

- Landings are used rather than catch. Cull-size and discarding has increased over time, and any changes in fishing practices over time could have caused changes in discard mortality. Also, the more recent use of 4-inch rings is expected to reduce discarding. Hence the ratio of landings to survey biomass may give a biased picture of trends in F.
- The fishing mortality from the 2-bin method will be sensitive to the length-ranges used for the two bins. The length slicing may induce a year-effect in the indices for the two size ranges, which may be partly responsible for the negative F estimates in more recent years on the Georges Bank, as abundance has increased.
- A tendency for relatively larger numbers of 80-100mm scallops compared with 100mm+ in the survey of the Georges Bank up to the mid 1990s, with the opposite trend in most subsequent years, will have weighted the 2-bin F estimates towards larger scallops in the second period, and vice versa in the earlier period.
- The difference between the annual estimates of F from the catch-biomass and the 2-bin methods vary non-randomly over time, probably due to the problems highlighted in the previous two bullet points. The ratio of the means of the two series will therefore differ in different parts of the series, and will drift non-randomly as new data are added. This suggests that re-scaling the catch:biomass estimates of F may lead to biased estimates of absolute F.
- The ratio of the two mean Fs will be biased anyway, as it is the ratio of two values with their own error distributions.

- It is not clear how the combination of data from closed and open areas will have affected the analysis
- The very low Fs from the Georges Bank catch-biomass ratio since the late 1990s reflect small landings relative to the survey biomass index. As discussed above, this could be affected by the change in discard rates depending upon discard mortality. Notably, the F estimates from the 2-bin method in the last three years are much larger than the re-scaled catch:biomass estimates, both in absolute terms and relative to the large values in the late 1980s and early 1990s.

It is concluded that the method currently used to estimate fishing mortality relative to the threshold value is likely to give rather inaccurate estimates of fishing mortality for recent years on the Georges Bank. Estimates from the catch:biomass and 2-bin methods for Mid Atlantic Bight scallops data appear more consistent over time.

3.2.7 Results of the assessment

Conclusions regarding trends in abundance from surveys are likely to be robust.

There are much greater problems in evaluating the status of the stock(s) in terms of whether overfishing is occurring, i.e. if fishing mortality is above the threshold. This is because the method for estimating F is probably not robust. Furthermore, it does not appear to make sense to derive a single estimate of fishing mortality for two areas with different trends in F and for which there are substantial closed areas with effectively zero fishing mortality. The conclusion of the working group that "overfishing is occurring" may still be appropriate, as fishing mortality on Georges Bank in recent years could be under-estimated.

3.2.8 Biomass and fishing mortality reference points

The fishing mortality reference points were updated to allow for recent changes in gear selectivity, using the length-based yield-per-recruit model. F_{MAX} is adopted as a proxy for F_{MSY}, as in the fish assessments reviewed at SARC39. The assessment scientists should plot the YPR curve and indicate F_{MAX} and $F_{0.1}$. Without such a plot it is not possible to know if the YPR curve is relatively flat at the F_{MAX} point, i.e. if this point is well defined in a domed YPR curve, and likely to be not far from $F_{MSY.}$

Calculation of B_{MAX} (proxy for B_{MSY}) as median recruitment from surveys multiplied by biomass per recruit at F_{MAX}, and comparing this with the current biomass index for the entire assessment area, has some problems:

- Recruitment for calculating YPR appears to be numbers at 40mm shell height. The surveys give a point estimate of numbers in the recruiting age class, representing a length at age distribution reflecting birth-date distribution and variability in growth between individuals. This distribution is not fully quantified – a length "slice" appears to be used. Unless the measure of recruitment is the same in both cases, an incorrect estimate of B_{MAX} will be obtained.

- The biomass index at any time will represent the sum of the biomass indices for all the year classes present in the population. The average biomass index over the long term will be equivalent to the sum of the biomass at successive ages in an "average" year class. If the recruitment distribution is skewed (e.g. lognormal), the correct "equilibrium" B_{MAX} will be the arithmetic mean from the distribution of year classes, not the geometric mean or median. This is a problem, for example, in Monte Carlo simulations of fish populations where stock recruit curves have been fitted assuming lognormal residual error whilst omitting to bias-correct the stock-recruit parameters.
- Consideration needs to be given to how you define "average" recruitment for the entire mosaic of scallop grounds off the Northeast coast, when there have been regional differences in productivity and several areas closed to fishing for part of the time period studied.

3.2.9 The proposed CASA method

The Panel was given an excellent presentation on the development of a Catch At Size Analysis (CASA) implementation for scallops. This method addressed a major criticism that the scallop assessment was not model-based. The use of a size-transition matrix to model length compositions is an established method (Sullivan *et al*. 1990; Punt *et al*. 1997), although not widely used. The Sea Scallop implementation takes a stock-synthesis approach and fits a wide variety of length compositions and other variables from surveys and the fishery, whilst keeping parameter numbers within reasonable bounds. The application benefits from having well-defined selectivity parameters for the survey dredge gear.

Further development of this approach is encouraged, and could indeed be extended to fish stocks such as bluefish where historic age compositions are questionable. Parallel with this development should be more extensive simulation testing of the method. Given the doubts over the accuracy of commercial fishery data and historic LFDs, which is reflected in a poor fit of the CASA to some fishery LFDs, a restricted application to survey-only data could also be tried for estimating relative biomass trends and mortality rates. This may allow application to smaller geographic units, which is not possible using fishery data because of inadequate LFD sampling at such small scales.

Sensitivity of the results to changes in growth rate and hence the size-transition matrix, or to spatial differences in growth, should be evaluated.

3.2.10 Recommendations of previous review

The SAW 32 workshop in April 2001 made the following recommendations: 1) adequacy of survey stratification design for separately monitoring open and closed areas should be investigated; 2) further depletion studies should be conducted using coincident dredging/photographic experiments in areas with different bottom types, scallop densities and size ranges. Multi-beam sonar or other methods should be used to map bottom types so that bottom type specific efficiencies can be used in the surveys; 3) length based yield-per-recruit should be further developed, accompanied by further work on equivalent length-based assessment methods; 4) updated information on scallop growth is needed for the assessment and calculation of reference points; 5) take advantage of opportunities afforded by closed areas to refine estimates of natural mortality and growth, and to estimate non-yield mortality from fishing.

The SAW39 assessment report discusses post-stratification schemes including a new scheme for the Nantucket Lightship closed area, but does not really address the adequacy of the current survey design for separate monitoring of open and closed areas.

Appendix 3 in the SARC 39 report presents a study of dredge efficiency using SMAST video data, but did not consider how bottom type affects efficiency. The issue is partly addressed in Appendix 2 of the SARC 39 report, but this looks specifically at the effect of rock chains in dredges deployed on different bottom types.

Length-based yield per recruit was applied this year, and the sensitivity of the results to a wide range of assumptions regarding growth and mortality was examined. The CASA model may be considered as compatible with length-based YPR, provided both models are structured to ensure equivalence.

Recommendations 4) and 5) were not addressed in SAW39.

3.2.11 Recommendations for future SeaScallop assessments

- The method used in this year's and the previous assessment for estimating fishing mortality is *ad-hoc* and may not be robust. The assessment should move to a model-based approach using observation-error models configured with regard to the nature of errors in variables. Further development of CASA is recommended, together with simulation studies to determine sensitivity and weaknesses. Estimation of biomass and fishing mortality reference points within the length-structured model is recommended to ensure compatibility (e.g. re-estimation of reference points and current status in each bootstrap run).

- More parsimonious models should also be explored. The similarity in general trends in fishing mortality from the rescaled catch:biomass ratios and from CASA (SAW 39 report Appendix Fig. 4-15) suggests that much of the dynamics is captured in the basic catch and survey data. The more complex CASA model uses a wide range of data simultaneously, with more appropriate estimation procedures, but the trade-off is an element of smoothing and possibly bias by assuming a constant size-transition matrix and natural mortality as well as integrating across large sea areas. In addition, some of the commercial fishery LFDs are based on limited sampling. Simple production models (e.g. ASPIC), or extensions including information on recruitment, could prove useful if applied at the scale of open and closed areas, if commercial catch data can also be extracted at this scale. Depletion of stocks followed by recovery in closures could allow good estimates of production model parameters along the lines of "adaptive probing experiments". This would have the advantage of providing standard reference points such as F_{MSY} and B_{MSY} for each area. Simple models also facilitate application of simulation models to explore performance of harvest-control rules including rotating closures. Such modelling becomes very complex if the assessment process is to be simulated and the assessments are using more complex models such as CASA.

- Given the potential for variation in growth to affect length-based assessments and potential yield, and for the productivity of different regions to vary with time due to changes in environment, scallop assessment reports should also present information about any relevant changes in regional habitat conditions. This includes changes in bottom temperatures, plankton production, larval drift patterns, seabed characteristics and densities of starfish and other predators.

- The present requirement to give assessment results and biological reference points for the entire Northeast area does not make sense given the regional variations in stock trends and productivity. For example, the entire stock complex may be classified as experiencing overfishing when in fact there is overfishing in the Mid Atlantic Bight but low exploitation rates on Georges Bank. Vessels exerting sustainable rates of exploitation in one sea area may then become subject to further management restrictions because of overfishing in a different sea area. It may be more appropriate to specify reference points and management approaches that can be applied regionally.

3.3 Atlantic Bluefish

3.3.1 Summary

The ASPIC biomass dynamic model presented at SARC39 cannot be accepted as a basis for fishery management for the following reasons:

- The recreational CPUE series contains a severe bias due to incorrect handling of the live-release data

- The NEFSC data used as an index of fishable biomass represents only 0-gp and 1-gp bluefish

- Residuals in the commercial CPUE data show strong autocorrelation indicating model mis-specification.

These problems were identified at an early stage in the presentation of the assessment, and are discussed in detail below.

3.3.2 Comparison with previous assessment (SAW 23)

The SAW 23 assessment of bluefish in March 1997 used a catch-at-age analysis (Integrated catch at Age Analysis, or ICA) to reconstruct historical spawning biomass, recruitment and fishing mortality up to 1995. The analysis was tuned with age-structured catch-per-unit-effort (CPUE) from recreational fisheries, and a range of age-structured and age-aggregated survey indices of population size. The assessment was rejected because of perceived problems with the accuracy of age composition data, including potential misallocation of ages using scales during earlier years.

The ICA model indicated a domed selection pattern across ages, and the resultant small values of fishing mortality on mature bluefish implied very large biomass values historically relative to landings. This contrasts markedly with the biomass dynamic model (ASPIC) results presented at SARC39, which gave higher fishing mortality rates and much lower biomass during the period 1982 - 1995 (20kt – 100kt compared with 120kt – 320 kt from ICA). Both methods show a depletion of biomass between the mid 1980s and the mid 1990s, by a factor of four using ASPIC and a factor of around 2.5 using ICA.

3.3.3 Fishery data used for the assessment

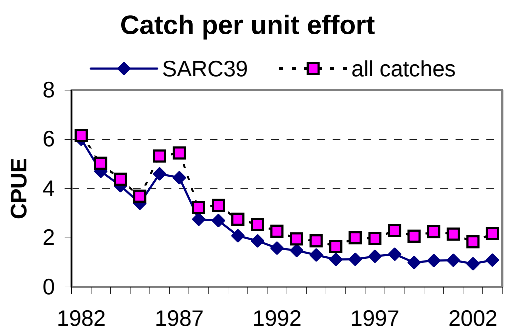
The Working Group report provides a commendable and very useful summary of the biology of the stock, the commercial and recreational fisheries, and the fishery and survey data. I would suggest that in future much of the information of this nature is placed in an annex rather than the body of the report. Important information missing from the report was the maturity ogives for males and females.

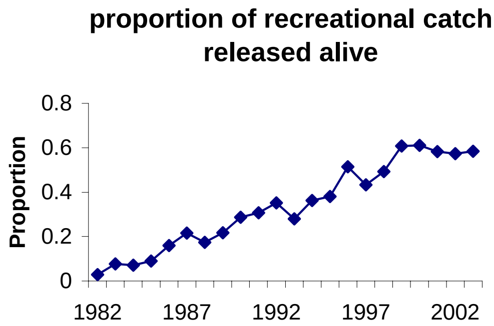
The commercial landings data were considered reliable, although these are only a relatively small fraction of the total catch. Landings increased four-fold from 1950 to the early 1980s and then declined by a factor of two during the following two decades. Unfortunately there are no reliable effort statistics to allow computation of LPUE data and hence to evaluate if the small catches in the 1950s reflected low abundance or small fishing mortality exerted on a large stock. No time-series of discards data are available. The SAW39 report indicates that discard rates in the commercial fishery are relatively low.

The recreational catch is estimated by telephonic and intercept surveys, and therefore has an associated sampling error that was not dealt with explicitly in the assessment model. This error will be replicated in the annual catch and the CPUE data. Information provided during discussion indicated coefficients of variation of 4 – 20% on estimated recreational catches by weight. The survey method appears to have been consistent over time since 1982.

The proportion of the recreational catch released alive has increased almost linearly from less than 10% up to 1985 to around 60% from 1999 (see text figure below). To calculate the total deaths due to fishing, it was assumed that 15% of recreational discards die. Given the large proportion of bluefish now being released alive, the discard mortality is a critical parameter yet is poorly known and should be investigated further.

The recreational CPUE data are a key data set for establishing trends in abundance of the stock. The data used in the SAW39 assessment were derived from trips that targeted bluefish (those in which bluefish were the primary or secondary target species). It was discovered that the CPUE had been calculated from the recreational fishery deaths rather than the total catch, by using the landings plus 15% of the live releases. The difference in trends for the two CPUE series is substantial:





Use of the recreational fishery deaths results in a steeper decline in CPUE since the 1980s. This has sufficient implications for the SAW39 assessment to invalidate the ASPIC model results.

The baseline ASPIC model was run assuming constant catchability over time in the recreational CPUE. There are many *a priori* reasons to expect this assumption to be violated over a period of over 20 years during which the abundance of bluefish has declined, angling technology has improved and there have been developments such as the moratorium on striped bass fishing that affect patterns of angling. As angler CPUE data are likely to remain an important source of information on stock trends for future assessments, data and methods of analysis should be sought to minimise potential changes in catchability**.** This may involve modifying the type of information collected during surveys to allow more appropriate stratification or standardization of the data, and tracking the catch-rates of individuals or groups of individuals over time in addition to random point estimates. Records of catch rates and specimen fish kept by angling clubs, or catch rates during regular competitions, may provide useful information, and may extend much further back in time than 1982.

3.3.4 Fishery independent data used for the assessment

The SAW39 report lists 12 separate surveys, mostly of juvenile bluefish, ranging from very localised surveys to some covering large parts of the stock's range. Poor coherence of temporal trends between many of the small-scale surveys may merely reflect inter-annual variability in distribution of young bluefish, and is not a reason to discard the information from these surveys. The two largest-scale surveys (NEFSC and SEAMAP), covering the northeast and southeast inshore regions, although only weakly correlated (R^2=0.3) show the same directional changes in CPUE from year to year in 11 out of 12 years, suggesting some large-scale coherence. It is recommended that methods be explored for combined analysis of all suitable survey data for 0-gp and 1-gp bluefish to derive an improved recruitment time-series.

The SAW39 assessment used only the NEFSC fall survey data on the grounds that it covered the largest area with consistent survey design and gear. However, the use of the mean catch weight per tow as an index of fishable biomass in ASPIC was not appropriate as the survey catch rates are predominantly 0-gp and 1-gp fish. The use of the NEFSC catch-rates of recruits as an index of total stock biomass further invalidates the assessment provided to SARC39.

3.3.5 The SAW39 assessment procedure

The choice of a biomass dynamic model has been made in response to criticism of the quality of the age composition data used previously. Biomass dynamic models can work well if there is sufficient data contrast in stock size and fishing effort to allow estimation of the parameters *K*, *q* and *r*. The bluefish data span a period of substantial decline in catch and CPUE, and a recent period of around a decade when CPUE and catch have been relatively stable. Unfortunately, the period of increasing catch in the late 1970s does not have an associated CPUE series with full age composition – the fall survey covers this period but as discussed earlier reflects mainly recruitment. The largely downward trend in catch and CPUE since the 1980s is likely to result in a poor ability to obtain unbiased estimates for the different parameters of the production model. This may also explain the extreme sensitivity of the production model fit to tiny deviations in the starting parameter estimates. Partially age-structured methods such as the Collie-Sissenwine model, in which production due to recruitment is explicitly modelled, are likely to be more robust and should be investigated further until full age-structured assessments are possible.

3.3.6 Results of the assessment

Problems with the input data are reflected in the model fit. The recreational CPUE show strongly autocorrelated residuals, whilst the survey series show much larger but less correlated residuals with several extreme values. With both series having been given equal weight, the survey residuals exert strong leverage. Of particular concern is the indication of a recent increase in biomass in the fitted values which could result from the large positive residuals in the NEFSC data in 1999 and 2003 exerting undue leverage (SAW39 bluefish report Fig. C12). Large residuals in the NEFSC series in 1974, 1981, 1984 and 1989 may also have undue leverage in the model fit.

The Working Group is commended for exploring the utility of other production model forms and the sensitivity to changes in catchability. The assumption of constant catchability in recreational CPUE is likely to give the most optimistic view of the state of the stock unless there has been a significant increase in less efficient anglers over time, and must remain an issue of some concern that needs to be addressed externally to the model through a more comprehensive analysis of recreational catch data.

ASPIC model runs carried out in 2002, 2003 and SARC39 gave unstable estimates of virgin biomass (*K*) of 288kt in 2002, 400kt in 2003 and 358kt in

2004 although estimates of F_{MSY} were less variable. A formal retrospective analysis of trends in fishing mortality and biomass was not presented.

3.3.7 Biomass and fishing mortality reference points

The reference points specified in Amendment 1 to the bluefish FMP are derived from an ASPIC run carried out in 1998. Subsequent ASPIC runs have modified these values substantially:

Assessment year	B_{MSY}	F_{MSY}
1998	108	0.40
2002	144	0.30
2003	200	0.26
2004	179	0.28

The lack of stability in reference points is a further indicator of problems with the model.

3.3.8 Recommendations of previous review

Recommendations contained in SAW/SARC 23 included: 1) improve the biological sampling particularly from recreational fishery; 2) carry out age readings on archived material; 3) carry out a study of tag mortality and retention rates; and 4) testing of sensitivity of bluefish assessment to assumptions regarding age-varying M, levels of age-0 discards, and selection pattern.

The SAW39 assessment report states that substantial progress has been made in resolving age problems, and that stock-synthesis models have been run. However, no results were presented. There have been no further studies on tagging mortality and tag retention.

3.3.9 Recommendations for future bluefish assessments

- Mortality of bluefish released by anglers is a key parameter because of the large proportion now released alive, and should be the subject of a more detailed investigation. This should include effect of any potentially significant factors such as fish size, sex, method of capture and season.

- Recreational CPUE is likely to remain an important source of information on relative trends in overall stock size of bluefish. It is imperative that the data be collected in a way that allows analysis of changes in angler behaviour, composition, technology or other factors that influence both the statistical distribution of individual CPUE and any changes in catchability over time.

- There is a need for an integrated analysis of the many different research surveys for juvenile bluefish. These surveys cover different regions using different gear types and provide data on 0-gp and 1-gp bluefish. It is recommended that a workshop be convened to evaluate 1) the quality of the individual data sets, 2) the potential ability of the surveys to index bluefish abundance at each age in the areas surveyed; 3) coherence of trends in localised surveys with trends in the nearby stations of the larger-scale surveys; and 4) methods for standardizing and combining data from small scale intensive surveys with large-scale less spatially intensive surveys to give improved indices of recruitment. Such a workshop would require consolidation of raw survey data from the different surveys into common databases.

- Collection of sufficient age composition data to allow the continued development of fully age structured assessment models is recommended, particularly in view of the unusual selectivity patterns estimated from earlier catch-at-age analyses (SAW 23).

- Pending the ability to apply full age-structured methods, the use of partially age-structured methods is recommended to allow explicit incorporation of survey estimates for 0-gp and 1-gp bluefish to estimate the contribution of recruitment to annual production. This would, however, require the commercial fishery catch and the recreational catches and CPUE to be disaggregated into recruits and older fish. The effect of poor data on discards of young bluefish in the commercial fishery on such an analysis requires evaluation.

- Development of tagging studies as a means of estimating mortality, selectivity and movements is recommended.

4. REFERENCES

Punt, A. E., R.B. Kennedy & S. D. Frusher. 1997. Estimating the size-transition matrix for Tasmanian rock lobster, *Jasus edwardsii*. *Mar. Freshwater Res., 48, 981-992*.

Sullivan, P. J., H-L. Lai & V. F. Gallucci. 1990. A catch-at-length analysis that incorporates a stochastic model of growth. *Can. J. Fish. Aquat. Sci. 47, 184-198.*

Appendix 1: Panelists

Chair:

Dr Andrew I.L. Payne (Centre for Environment, Fisheries and Aquaculture Sciences, Lowestoft, UK)

Panel members:

Dr Michael J. Armstrong (Centre for Environment, Fisheries and Aquaculture Sciences, Lowestoft, UK)
Dr Din Chen (International Pacific Halibut Commission, University of Washington, Seattle, USA)
Dr Paul Medley (Consultant, Alne, UK)

Appendix 2: Terms of Reference

A. Black seabass

1. Characterize the commercial and recreational catch data (including length distributions).
2. Update Northeast Fisheries Science Center (NEFSC) survey indices and evaluate appropriate state survey indices.
3. Summarize tagging program results (NEFSC, Virginia, New Jersey).
4. Develop tag-based estimate(s) of exploitation.
5. Evaluate use of index-based methods for estimating relative Fs.
6. Re-evaluate biological reference points.

B. Sea scallop

1. Update status of the Georges Bank, Mid Atlantic Bight and Gulf of Maine sea scallop resources through 2003 using all applicable information fishery dependent information and fishery independent surveys (e.g. NEFSC trawl survey, SMAST video survey and others as appropriate). Provide estimates of fishing mortality and stock size. Characterize uncertainty in the estimates.
2. Evaluate stock status relative to current reference points.
3. Provide short-term projections of stock biomass and catches consistent with target fishing mortality rates.
4. Update estimates of biological reference points (e.g. B_{MSY}, F_{MSY}) using revised biological and fishery data, as appropriate.
5. Evaluate information provided by various current survey approaches and suggest possible ways to integrate their results.
6. Continue the development stock assessment modelling approaches that integrate all appropriate sources of fishery dependent and fishery-independent data.

C. Bluefish

1. Characterize the commercial and recreational catch, including landings and discards.
2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates.
3. Evaluate and either update or re-estimate biological reference points, as appropriate.
4. Where appropriate, estimate a TAC and/or TAL based on stock status and target mortality rate for the year following the terminal assessment year.
5. If stock assessments are possible,
    a. Provide short-term projections (2-3 years) of stock status under various TAC/F strategies, and
    b. Evaluate current and projected stock status against existing rebuilding and recovery schedules, as appropriate.

Appendix 3: Agenda

39TH NORTHEAST REGIONAL STOCK ASSESSMENT WORKSHOP (SAW39)
STOCK ASSESSMENT REVIEW COMMITTEE (SARC) MEETING

Aquarium Conference Room - Northeast Fisheries Science Center
Woods Hole, Massachusetts
7-10 June 2004

Date and Subject	Presenter	Panel lead	Rapporteur
MONDAY, 7 June (13:00 – 17:30)			
Welcome	**John Boreman**, Center Director		
Introduction	**Terry Smith**, SAW Chairman		
Agenda & Conduct of meeting	**Andy Payne,** SARC Chairman		
Black Sea Bass (A)	**Gary Shepherd**	**Din Chen**	**Laurel Col**
SARC Discussion	**Andy Payne**		
TUESDAY, 8 June (08:30 - 18:00)			
Sea Scallop (B)	**Dvora Hart**	**Paul Medley**	**Larry Jacobson**
SARC Discussion	**Andy Payne**		
WEDNESDAY, 9 June (09:00 - 17:00)			
Bluefish (C)	**Jessica Coakley**	**Mike Armstrong**	**Laura Lee**
SARC Discussion	**Andy Payne**		
THURSDAY, 10 June (09:00 - finish)			
Close discussion and report preparation			

Appendix 4: Bibliography

A1: **Assessment of the Northern Stock of Black Sea Bass**. Consensus Assessment Summary and Report of the Coastal/Pelagic Working Group Report prepared for the Stock Assessment Review Committee (SARC), May 4, 2004. 3 pp. + 82 pp.

A2: **H. Black Sea Bass**. *In* 27th Northeast Regional Stock Assessment Workshop (27th SAW). Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NEFSC Ref. Doc. 98-15.

B1: **Sea Scallop Assessment**. Summary and Report of the Invertebrate Subcommittee prepared for the Stock Assessment Review Committee (SARC), May 25, 2004. 3 pp. + 120 pp.

B2: **B. Sea Scallop Advisory Report**. *In* 32nd Northeast Regional Stock Assessment Workshop (32nd SAW). Public Review Workshop. NEFSC Ref. Doc. 01-04.

C1: **Bluefish Assessment**. Summary and Working Group Report of the ASFMC Bluefish Assessment Committee prepared for the 39th Stock assessment Review Committee (SARC), June 2004. 5 pp. + 63 pp.

C2: **C. Bluefish**. *In* 23rd Northeast Regional Stock Assessment Workshop (23rd SAW). Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NEFSC Ref. Doc. 97-05.

Appendix 5: Statement of Work

Consulting Agreement between the University of Miami and CEFAS, Dr. Michael Armstrong

May 13, 2004

General

The Northeast Regional Stock Assessment Review Committee meeting (SARC) is a formal, multiple day meeting of stock assessment experts who serve as a peer-review panel for several tabled stock assessments. The SARC is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes peer assessment development (SAW Working Groups or ASMFC technical committees), assessment peer review, public presentations, and document publication.

Designee will serve as a panellist on the 39th Stock Assessment Review Committee panel. The panel will convene at the Woods Hole Laboratory of the Northeast Fisheries Science Center in Woods Hole, Massachusetts, the week of 7 June 2004 (7-10 June) to review assessments for sea scallop (*Placopecten magellanicus*), black sea bass (*Centropristis striata*), and bluefish (*Pomatomus saltatrix*).

Specific

A panellist's duties will occupy a maximum of 14 workdays; a few days prior to the meeting for document review; the SARC meeting; and a few days following the meeting to prepare a Review Report. The SARC Review Report will be provided to the SARC chair who will produce a SARC Meeting Report summarizing the individual Review Reports.

Roles and responsibilities:

(1) Prior to the meeting: review the Working Group Reports.

(2) During the meeting: participate, as a peer, in panel discussions on assessment validity, results, recommendations, and conclusions especially with respect to the adequacy of the assessments reviewed in serving as a basis for providing scientific advice to management.

(3) After the meeting: prepare an individual Review Report which provides an executive summary, a review of activities and, for each stock assessment reviewed, a summary of findings and recommendations which emerge from the findings, all in the context of responsiveness to the Terms of Reference for each assessment.

(4) No later than June 25, 2004, submit a written report[1] consisting of the findings, analysis, and conclusions, addressed to the “University of Miami Independent System for Peer Review,” and sent to Dr. David Sampson, via e-mail to David.Sampson@oregonstate.edu and to Mr. Manoj Shivlani via e-mail to mshivlani@rsmas.miami.edu.

No consensus opinion between the CIE reviewers is sought and all SARC reports will be the product of the individual CIE reviewer or Chairperson.

Contact person:
Dr. Terrence P. Smith, NEFSC, Woods Hole, SAW Chairman, 508-495-2230, Terry.Smith@noaa.gov

[1] The written report will undergo an internal CIE review before it is considered final. After completion, the CIE will create a PDF version of the written report that will be submitted to NOAA Fisheries and the consultant.