
**REPORT ON THE 39TH NORTHEAST REGIONAL
STOCK ASSESSMENT WORKSHOP (SAW-39)
STOCK ASSESSMENT REVIEW COMMITTEE (SARC)
MEETING**

by

Andrew I.L. Payne
SARC-39 Chair (Appointed by CIE)

*CEFAS
Pakefield Road
Lowestoft
Suffolk NR33 0HT
United Kingdom*

www.cefasc.co.uk



Contents

- 1. Executive Summary and Recommendations 3
- 2. Background, preliminaries and documentation 4
- 3. Conduct of the meeting 5
- 4. Summary of Panellists' reports 7
 - A. Black seabass 7
 - B. Sea scallop 10
 - C. Bluefish 13
 - D. General recommendations 15
- 5. Final comments 16
- Appendix 1: Panellists 18
- Appendix 2: Terms of Reference 19
- Appendix 3: Agenda 20
- Appendix 4: Bibliography 21
- Appendix 5: Statement of Work 22

1. Executive Summary and Recommendations

Three stocks were covered in this review, black seabass, sea scallop and bluefish. In most cases, the terms of reference supplied to the panel were the same as those addressed by the presenting working group. In future, it is crucial that they all be the same. The established system of assigning a panel member and a local expert as stock leader and rapporteur (to produce a comprehensive *aide memoire*), respectively, was followed, worked well, and should be maintained. A day of closed discussions by the panel was scheduled in the agenda, and this facilitated the swift production of expert review and Chair reports through a closed discussion of interpretations. This procedure should be followed in future, but as much time as possible should be allowed between submission of the expert reviews to the Chair and to the CIE, and the submission of the Chairman's summary report. Comprehensive, accurate summarizing is only possible given sufficient time for consideration. The comments and summaries per stock are listed in text and here in no order of priority.

For **black seabass**, the assessment is acceptable as a basis for management advice, though the basis for the reference points was questioned. Stock health seems to be improving and, though little use is made of commercial fishery data in the assessments, excellent survey and tagging data are applied. The tagging programme itself is impressive, and the data lend themselves to future enhanced uptake in modelling and assessment. In terms of specific recommendations, the panel stressed the need for greater evaluation of regional survey data, adequate sampling of commercial and recreational data (specifically, discard rates are not well known), combining the survey indices through use of GLM or GAM, presentation of confidence limits, age determination of representative samples, construction of a population model, the application of more-advanced modelling techniques to the tagging data, efforts to raise awareness of the (value of the) tagging programme, and the need to seek evidence for environmental drivers of the catchability trends.

Sea scallop abundance is well estimated, but it is not that obvious that overfishing is occurring. Seemingly counter trends are evident for the Georges Bank and Mid Atlantic Bight, and the panel questioned the appropriateness of treating the two areas as a unit stock for assessment and management. Discard information is not good, and discard mortality needs to be evaluated better. Non-reporting of catches could be a problem, and growth rings in scallop shells could be a useful future source of information. Natural mortality is estimated well from clapper ratios, and the proposed application of the Catch-at-Size Analysis (CASA) model has merit and should proceed apace. Recommendations include moving the assessment procedure to a model-based approach, further development of the CASA model, including sensitivity testing, consideration of management through rotational areas of exploitation, evaluation of the relationship between shell height and meat weight, evaluation of the appropriateness of other models, studies on predation rates, collection of more habitat information, and, if feasible, regionalizing of the assessment and management.

The **bluefish** assessment presented was deemed inappropriate for management, and was rejected. Input data negatively influence the model fit, and there was therefore no concrete evidence that the stock was recovering. Indeed, some indications are that it is still at a low ebb. A recreational catch rate abundance index needs to be developed, and discard rates and mortality need to be better known. Calculated reference points seem to be highly variable year on year. Specific recommendations to put advice and management of this stock on a stronger footing include gaining clarity on release mortality rates, probing the questionable assumption of constant catchability, the use of both survey and commercial catch rates in future assessments, that declared catches be not simply assumed to be accurate, integrating the manifold different survey time-series, reducing fishing mortality, collection of age material, evaluation of assessment methodology for bluefish elsewhere in the world, and use of the tagging programme to evaluate mortality, selectivity, movements, and tag retention.

A few general (to all stocks) recommendations were also made, notably to find means of combining the manifold different abundance indices through GLM or GAM, to develop a checklist of standardised diagnostics output, to facilitate reviewing by assessment scientists and working groups, that assumptions and errors should be tested through simulation, where possible, that some sort of simple method (e.g. decision tables) be developed to allow managers to assimilate risks and uncertainty in the assessments, and that the sampling error in recreational catches should be included in assessment models.

2. Background, preliminaries and documentation

The panel met from 7 to 10 June 2004 in the Aquarium Room of the NEFSC Woods Hole Laboratory, with a Chair and 3 panellists, as listed in Appendix 1. The terms of reference of SARC-39 are outlined in Appendix 2, the final Agenda in Appendix 3, and the Bibliography consulted prior to and during the meeting in Appendix 4. The Chair's personal Statement of Work is listed in Appendix 5.

The documentation for the meeting in terms of draft agenda and assessment reports for review arrived electronically from Woods Hole on 1 June, and the SAW Chairman provided hard copy of the previous SARC assessments and evaluations for the same three stocks on arrival in Cape Cod on 5 June. All electronic material was provided in easily accessible form as either pdf or Word files. The NMFS contact for the SARC was Terry Smith (SAW Chair), who competently facilitated the distribution and circulation of documentation and was responsible for all "housekeeping", as well as ensuring a ready and appreciated supply of refreshments at the meeting. He also facilitated electronic links through a shared drive for panellists at the meeting, restricting the costly and time-consuming need to produce hard copy, though hard copy was available to those who requested it. On request too, Terry provided a

large locator map for the panellists, who were not so knowledgeable about US and NE Atlantic geography as the presenters and other participants.

Between 1 June and the commencement of the meeting, I familiarized myself with the documentation, specifically the methodology and assumptions inherent in each assessment. Terry Smith gave me valuable background on the standard meeting procedures and clear direction of what was expected as output from the meeting itself. He also engaged me in electronic discussion about the meeting agenda, specifically the order in which I wished to conduct discourse on the various stocks, facilitating the presence of the relevant staff and those interested in the debate at their times of availability. An evening meeting on 6 June in Woods Hole, at which final arrangements were made and clarifications given, followed up this electronic discussion.

Both Terry and I noted that there were some apparent anomalies with the terms of reference. Terry explained to me that the process by which the terms of reference were finalized was rigorous and well known. However, perhaps owing to his unavailability at the Woods Hole Laboratory for a significant period at the time when the working groups and subcommittees were meeting on the three stocks, the given terms of reference had apparently been “tweaked” by participants to meet their own interpretations of what was requested. The SARC-39 process is slightly changed from that of many previous SARCs in that the panel has been curtailed to a few international scientists only, and that the objective now is more to review the science and assessments than to attempt to come up with best management advice. This is a huge improvement on the previous process, but it was clear that both experienced panellists and presenters were still coming to terms with the new process. Also, as will be seen later, one of the processes I undertook during each review was to evaluate the extent and accuracy with which each term of reference had been met. Working Groups and Subcommittees, having interpreted the priority of their (tweaked) terms of reference in their own ways, therefore paid different levels of attention to the various aims. This situation became obvious during each presentation, and I strongly urge future presenters to SARCs to be more rigorous in following their agreed terms of reference. Only the agreed terms of reference for each stock are provided to the panellists in advance of the meeting.

3. Conduct of the meeting

The meeting was convened at 13:00 on 7 June. Participants were warmly welcomed by the NEFSC Director, John Boreman, before the SAW Chair officially opened the meeting by introducing the panellists, all appointed by the CIE. He then handed the meeting over to me and I explained what I wanted to achieve for each stock (as per the Terms of Reference - Appendix 2). Specifically, I stressed that I would seek to determine the extent to which each assessment working group had met the terms of reference they had been given, and if they had fallen short on certain of these, would likely recommend closer adherence intersessionally (before the stock was subject to a future SARC evaluation). Although the SARC process had been amended as from

this meeting, I felt also that there was merit in determining whether or not the previous SARC's research recommendations had been met; again, if they had not, I would try to determine why not and, if appropriate, add them to the recommendations to be addressed intersessionally.

After the preliminaries of personal identification by panel members, the agenda was confirmed and the order of debate stayed the same as initially agreed. Thus, the meeting commenced with a presentation on black seabass by Gary Shepherd (Paper A1), followed by in-depth discussion by panellists. The presentation was usefully sectioned so that clarity and comments could be sought and made throughout, by both panellists and visitors from the floor. The same process was followed on the second and third days respectively for sea scallop (presenter Dvora Hart, assisted by Larry Jacobson; Paper B1) and bluefish (presenter Jessica Coakley; Paper C1).

For each stock, one panellist was designated as SARC leader, to liaise with myself as SARC Chair, the presenters, and the rapporteurs in ensuring that the issues raised and the targeted output were achieved to time and quality. The rapporteurs were nominated in advance by the presenters for each stock, though their report was not aimed at publication. The SARC leaders and rapporteurs were respectively Din Chen and Laurel Col (black seabass), Paul Medley and Larry Jacobson (sea scallop), and Mike Armstrong and Laura Lee (bluefish). This system worked very well and allowed me as Chair to concentrate solely on whatever issue was on the table in the knowledge that production of *aides memoire* was in capable hands and that all three such documents would be available for panellists before they left the meeting.

A small part of the Wednesday afternoon and the whole of the final day of the SARC (Thursday) were devoted to a closed discussion by the SARC panel on the material presented and to ensuring that the interpretation of each panel member was consistent. This debate was entered into not to seek consensus on output and report content, but to ensure that the basis on which the panellists' individual comments and recommendations were being made was the same. Further, one or two runs of data were made while we were all still present in Woods Hole, to seek clarity of some of the trends underlying the analyses and assessments of status. I was fully satisfied with the manner in which this latter part of the meeting was conducted. Indeed, the day set aside for these panel discussions in Woods Hole facilitated our tasks of report writing and submission of reports to the CIE.

I adjourned the meeting of the SARC-39 panel on Thursday afternoon.

I feel that one point relating to the SARC process deserves specific mention here. The new process seemingly requires the contracting of a number of practicing experts in stock assessment and of a chair conversant with the techniques, but not necessarily as technically astute in the detailed analyses. The duties of the chair are indeed clearly specified in the Statement of Work I agreed to before the meeting, namely to become conversant with the material presented, to ensure smooth running of the meeting, and to summarize the findings and recommendations of the panellists for consideration of the

customer. This I believe I achieved, but I have to stress that the latter was only possible through the agreement while at Woods Hole of all panellists to deadlines that allowed me to meet my own tight deadline for completion (2 July). Deadlines have to be tight, but international reviewers by their very description tend to have broad commitments that can upset the best-laid plans for producing a summary report. Simply, the longer the time between panellists' submission to the chair and to the CIE of their draft reports and the absolute deadline for the chair to submit his final report, the better.

Finally, in terms of my own limitations regarding up-to-date experience of stock assessment technology and of my knowledge of the SARC process itself, I record my gratitude specifically to Paul Rago and Terry Smith for the support they provided me at the meeting. Paul was always on hand to advise, and Terry went out of his way to ensure that chairing the meeting was indeed a pleasure and, hopefully, a success.

4. Summary of Panellists' reports

The reviews presented by the Panellists were, to my mind, accurate and comprehensive, and all three addressed their contractual obligations to the letter. In writing this summary, I did not seek any form of consensus, but readers will be aware from the detail in the individual reports that there was not only consensus but also unanimity in many issues about the three stocks. To summarize the three experts' reviews in a manner suitable for the NMFS to act, I have highlighted the main points of all in a single overview without attributing them to any reviewer. The summary is constructed in bullet-point fashion, without any attempt at prioritisation, to ensure that all main points are made. Recommendations are similarly dealt with, though there are also some points that can be construed as "recommendations" in the main bullet-point summary.

I have arranged my summary by species to facilitate future action by NMFS and presentation to the Fishery Management Council. All issues (facts, interpretations, recommendations) about which the reviewers felt most strongly were identified to me in the draft reviews, and they have been incorporated. However, it must be stressed that any such summary of several reports that is written by a single different person will contain some degree of personal subjectivity and interpretation, as well as sometimes being weighted towards the wording used by one or two rather than all the reviewers. Nevertheless, I have attempted to be as comprehensive as possible in what follows, though it is always possible that I may have inadvertently omitted some issues deserving of mention in one or other of the reports. Full and comprehensive (often technical) detail is available in the three separate reviews, which should be read in tandem with this summary.

A. Black seabass

- The assessment presented at SARC-39 is acceptable for fishery management because (1) the Spring NEFSC survey index of biomass for

fish ≥ 22 cm remains well above the threshold of $0.5 \times$ 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; (2) considerable research effort has been devoted to carrying out a well-designed and implemented tagging programme, and the results indicate that recent rates of fishing mortality are below the currently adopted F_{MSY} proxy ($F_{max} = 0.33$ from yield per recruit).

- The biomass reference point 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, it 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.
- There are other indicators that stock health is improving. Analyses confirm that recent fishing mortality is less than required for the stock to replace itself, so further growth of biomass remains possible. Moreover, an apparent progressive shift in the composition of landings towards “large” and “jumbo” categories since the late 1990s suggests the influence of a strong recruitment to the fishery, although that trend could be the result of “high grading”.
- Little use is made of commercial fishery data in the assessment, but the accuracy of commercial landings data has improved over time as a consequence of changes in the reporting system. However, there is seemingly no information on discarding. Recreational catches are at a level similar to commercial landings, but are highly variable from year to year, perhaps reflecting inaccuracies inherent in the survey method for estimating recreational landings and discards.
- Numerous research survey series of data are available, some limited in spatial extent, others covering a larger part of the eastern US coast. Winter-spring surveys consistently show an increase in abundance since the late 1990s, although with different years of peak abundance, and in most cases a sharp decline after the recent peak. Despite the multiplicity of surveys, only one (NEFSC Spring) is used in the assessment, because it covers a large area with consistent methods over time.
- A potential area of concern is that cross-shelf migration of black seabass could result in the spring survey catch rates being greatest at the outer boundary of the survey grid. If the stock distribution extended beyond the survey to differing extents in different years, or if the vertical distribution along the shelf edge changed in response to environmental conditions, strong year-effects would be apparent. Without age composition data, it is difficult to evaluate whether the periodic occurrence of large catch rates (as apparent since the late 1990s) is an effect of survey catchability or strong recruitment.
- A comprehensive tagging study has been initiated in response to the recommendations of the previous SARC. One very important output from this work will be enhanced knowledge of the variability in mortality estimates and of the uncertainties around 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.

- A Petersen model is not appropriate for modelling the outputs of the tagging programme, given the different fishing mortalities between regions. Consequently, there could be value in attempting to apply a more advanced model, such as that of Brownie with its migration extension, to estimate migration patterns among the different regions. Also, it may be possible to develop a selectivity function from the data on tag release at length.
- Overall, in terms of assessment, the dependence of management decisions on raw, age-aggregated survey indices is a shortcoming; there is no way to distinguish year-effects in catchability from genuine changes in biomass. Without robust data on year-class variation, those are difficult to forecast, particularly if they are driven by recruitment attributable to a truncated age composition.

Recommendations

Of the terms of reference set for black seabass, all were met except for the evaluation of biological reference points. However, their estimation without the establishment of a population model incorporating growth and mortality was not appropriate. In addition, of the research recommendations made the previous time black seabass were subject to a SARC evaluation, apparently only the single one dealing with the establishment of a tagging programme was taken up. The others were either mildly considered or apparently overlooked, perhaps for good reason. Those recommendations have therefore been overtaken by the current ones, which follow in no specific order of priority.

- × More comprehensive evaluation of regional survey data is required to give more integrated indices of recruitment. For example, catch rates of recruits can be modelled as a function of location, time of year and gear type in the surveys, to provide standardized indices. Good understanding of recent recruitment dynamics is essential for forecasting; the stock appears to have variable recruitment, and periodic population growth may be due to good recruitments.
- × Adequate sampling of both commercial and recreational catches should be implemented with a view to improving knowledge of discarding and what affects it, so reducing one of the uncertainties inherent in the catch series.
- × Both accuracy and completeness of catch data, particularly recreational catch, should be investigated to explain the unusual interannual variability.
- × Attempts should be made to extract as much information as possible from all time-series considered appropriate using, for example, a GLM or GAM approach to combine the various surveys and gear types into a standardized index.
- × Confidence limits for survey-based estimates of recreational catch should be derived and presented.
- × Ageing of samples of black seabass should be initiated as soon as possible, and survey indices need to be disaggregated by age to identify the impact of year-class variation in the biomass index and to investigate the magnitude of year effects.

- × A standard assessment based on a population model should be developed for the stock. A catch-at-age model would seem to be most appropriate
- × Clarification is needed on whether the bias introduced on back-transforming from the length-weight relationship has been corrected for in the assessment. If not, it should be.
- × If financially feasible, tagging studies should continue (at least sporadically), to permit return rates over longer periods and the stability of estimates of exploitation rate to be established. Further, long-term data on rates of tag loss need to be collected through the tagging programme.
- × A more sophisticated analytical model such as that of Brownie with a migration extension should be applied to the tagging data.
- × Improved education and awareness programmes should be initiated in an attempt to improve tag return rates. Those who do not return tags, regardless of the reward, cannot understand how and why the programme is being carried out. One way to promote awareness is to involve fishers more in the tagging programme.
- × 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.

B. Sea scallop

- Trends in sea scallop abundance are well estimated from the intensive dredge surveys, more so than at previous SARC evaluations. However, there are 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, 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 many closed areas with effectively zero fishing mortality. However, there may still be overfishing, because fishing mortality on the Georges Bank in recent years may be underestimated in the assessment.
- 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. Fishing mortality is less well estimated, and although 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.
- Rotational closures are worth considering as a form of management. As biomass recovery is largely attributable to closed areas, spatial management would appear to be a good option.
- It will be necessary in future to address recruitment from Canada. It may be possible to use indices of spawning stock biomass and recruitment from the Canadian Georges Bank.
- There are differences in the historical survey estimates between this SARC evaluation and the previous one, presumably the consequence of re-working indices. 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 two- to threefold decline over time.

- The catch/landings data used for calculation of F 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. Non-reporting of catches causes underestimates of fishing mortality using the catch : biomass method.
- Fishery discards affect the catch data used in both the estimates of fishing mortality and the proposed CASA method. Observer data show an increase in cull size and proportion discarded since the early 1990s. It is assumed that 20% of the discards die, but this may vary considerably, depending on ambient fishing conditions. This assumed rate could well be too high, given that fishers now utilize 4-inch rings to avoid catching small scallops, thus reducing discarding.
- At present, growth rings in scallop shells cannot be identified reliably, but when they can, transition matrices can be empirically based, using methods similar to age-length keys. In the current form of the model, growth rings are only used to estimate variations in growth. Also, non-parametric density estimation should prove useful in modelling growth increment probability, so reducing the assumptions based on the use of the growth model and gamma distribution for growth variability.
- There may be problems and biases with length frequency distributions prior to 1994, and those for the period 1985-1993 are excluded from the CASA model. Only seagoing samples have been used for the more recent years, but in some years the number of trips sampled has been very small.
- The sources of error in the fishery catch and length 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.
- The method of estimating natural mortality using clapper ratios is novel and useful. A value of M of 0.1 is appropriate for use in assessments for scallops >40 mm shell height.
- There is great merit in continuing the implementation of a CASA model for scallops, so addressing the major criticism that the scallop assessment is not model-based. 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, while keeping parameter numbers within reasonable bounds. The application also benefits from having well-defined selectivity parameters for the survey dredge gear.

Recommendations

Of the terms of reference set for the stock, most were addressed adequately. However, several recommendations from the previous SARC at which this stock was addressed were not followed, and others only paid scant attention to. In particular, it was disappointing that greater progress had not been made towards use of a population model to assess the state of the stock. Much of the current presentation highlighted the proposed new application to the stock

of the CASA model, and the outputs from this work likely overtake the previous recommendations. Consequently, the main recommendations for this stock follow below.

- × 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 and its application to separate management units (as a minimum, the Mid Atlantic Bight and Georges Bank) is recommended. 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).
- × If rotational areas are being considered, zones could be managed separately, as long as joint spawning stock biomass is maintained.
- × Sensitivity of the CASA results to changes in growth rate and hence the size-transition matrix, or to spatial differences in growth, should be evaluated.
- × A YPR curve should be provided in future, indicating F_{MAX} and $F_{0.1}$.
- × The relationship between shell height and meat weight is crucial in the assessments and should be robust.
- × Models other than CASA need to be explored. The similarity in general trends in fishing mortality from the re-scaled catch : biomass ratios and from CASA 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.
- × 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, provided commercial catch data can be extracted at this scale. This would have the advantage of providing standard reference points such as F_{MSY} and B_{MSY} for each area, and facilitate application of simulation models to explore performance of harvest-control rules, including rotational closures.
- × 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 as a consequence of changes in the environment, scallop assessment reports should present information about any changes in regional habitat conditions (e.g. bottom temperature, plankton production, larval drift patterns, seabed characteristics, densities of starfish and other predators).
- × Studies on predation rates in areas covered by the dredge surveys may provide useful information to help interpret apparent changes in abundance. Further, in view of the move towards length-based assessments that require growth parameters, a better understanding of spatial and temporal variations in both the mean and the variance of growth rates, and the causes of such variation, should be sought.

C. Bluefish

- The ASPIC biomass dynamic model presented at SARC-39 should not be accepted as a basis for fishery management because (i) the recreational catch rate series contains a severe bias attributable to incorrect handling of the live-release data, (ii) the NEFSC data used as an index of fishable biomass represent only 0- and 1-group bluefish, (iii) residuals in the commercial catch rate data show strong autocorrelation, indicating model mis-specification, (iv) the model is too sensitive for the population growth parameter r .
- Estimates of fishing mortality and total stock biomass for the current year are highly uncertain.
- There is no real evidence that the stock is recovering, and the fact that independent population indices give broadly the same signal (substantial decline since the beginning of the series) indicates that the stock may well be below 50% of its unexploited state (below B_{MSY}).
- Problems with the input data are reflected in the model fit. The recreational catch rates show strongly autocorrelated residuals; the survey series show much larger but less correlated residuals, with several extreme values. Catch uncertainty ought to be taken into account in future stock assessments.
- Useful information is available on the biology of the stock, the commercial and recreational fisheries, and the survey. In particular, the commercial landings data seem to be reliable, although they represent only a relatively small fraction of the total catch; they indicate little discarding.
- The recreational catch is estimated by telephone and intercept surveys, and therefore has an associated sampling error that was not dealt with explicitly in the assessment model. The survey method appears to have been consistent since 1982.
- The survey index suffers from problems such as its lack of coverage of the whole stock, the fact that it is not directed just at bluefish, and that catches tend to be smaller bluefish only.
- The proportion of the recreational catch released alive has increased almost linearly from <10% up to 1985 to some 60% from 1999. To calculate the total deaths attributable to fishing, it is assumed that 15% of recreational discards die. However, given the large proportion of bluefish now released alive, discard mortality is a critical parameter, yet remains poorly known.
- The bluefish data span a period of substantial decline in catch and catch rate, and a recent period of around a decade when both parameters have been relatively stable. Unfortunately, the period of increasing catch in the late 1970s does not have an associated series of catch rate with full age composition; the fall survey covers this period, but reflects mainly recruitment. The largely downward trend in catch and catch rate since the 1980s is likely to result in a poor ability to obtain unbiased estimates of the different parameters of the production model.
- The reference points specified in Amendment 1 to the bluefish Fishery Management Plan are derived from an ASPIC run carried out in 1998.

Subsequent ASPIC runs have modified these values substantially, and the lack of stability in reference points is a further indicator of problems with the model.

Recommendations

For this stock, there were anomalies between the terms of reference given to the panel and those addressed by the working group presenting the assessment, so although it was clear that many had been met, it was equally obvious that some had not. Further, in evaluating progress against the recommendations of the previous SARC at which bluefish were evaluated, it was not clear what had indeed been accomplished, because although it was said that ageing problems had been resolved and stock-synthesis models run, no results were presented. Likewise, there did not seem to have been any further work on tag retention or tagging mortality. Recommendations for future work include some of these omissions and are listed below.

- × The 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 catch rate is important, so the data should be collected in a manner that allows analysis of changes in angler behaviour, composition, technology, or other factors that influence both the statistical distribution of individual catch rate and changes in catchability over time.
- × An assumption of constant catchability in recreational catch rates is likely to give an 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 concern that needs to be addressed externally to the model, through a more comprehensive analysis of recreational catch data.
- × Catch rate and survey indices should both continue to be used for assessment purposes, if possible. However, models other than a catch rate index should at least be considered.
- × Terceiro (2003, Fishery Bulletin 101, pp. 653-672) has done much of the groundwork needed to develop a recreational catch rate abundance index. Poisson quasi-likelihood may be the simplest error model to apply. If possible, all trips should be used, and targeting should be allowed for as factor in the GLM.
- × Catches should not be presumed to be exact, but can be fitted through some likelihood function for discrepancies between observed and estimated catch in the population model. The likelihood can use the standard error of the catch estimate.
- × There is a need for an integrated analysis of the many different research surveys for juvenile bluefish. The surveys cover different regions using different gear types and provide data on 0- and 1-group bluefish. It is recommended that serious consideration be given to convening a workshop to evaluate: 1) the quality of the individual data sets; 2) the potential ability of the surveys to index bluefish abundance at age in the areas surveyed; 3) coherence of trends in localized surveys with trends in 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.

- × Care should be taken when using a GLM index approach that information relevant to changes in stock size is not mistakenly removed. A better approach might be to integrate the GLM into a population model.
- × Reducing fishing mortality to allow the abundance indices to increase could provide useful information on the productivity of the stock. A much-improved assessment may be obtained when a recovery has taken place.
- × Age composition data should be collected to allow continued development of fully age-structured assessment models, particularly in light of the unusual selectivity patterns estimated from earlier catch-at-age analyses.
- × Stock assessment methods applied to bluefish elsewhere in the world should be evaluated for applicability to the NE US situation.
- × Pending ability to apply full age-structured methods, the use of partially age-structured methods such as the Collie-Sissenwine model is recommended to allow explicit incorporation of survey estimates for 0- and 1-group fish, so estimating the contribution of recruitment to annual production. This would require that the commercial fishery and recreational catches and cpue 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.
- × Global search algorithms (e.g. genetic algorithms) should be used for parameters if an ASPIC model is used in future.
- × Maturity ogives need to be constructed and presented in future assessments.
- × As the current assessment has been rejected, and the status of the stock is unknown, the total allowable landings specification should continue at current value.
- × The feasibility of using tagging studies to estimate mortality, selectivity and movements, as well as to determine tag retention, should be investigated.

D. General recommendations

Some general recommendations common to one or more of the assessments evaluated above, some already listed and some not, are given below.

- × It was clear to the panel that, for at least black seabass and bluefish, and likely other stocks too, some data series were not being included in evaluations of stock status. It was therefore recommended strongly that attempts be made to extract as much information as possible from all series considered appropriate for each stock using, for example, a GLM or GAM approach to combine the various surveys and gear types into a standardized index. This objective could be initiated through convening a workshop at which State and Federal scientists could debate many such data series and the appropriateness and ways of combining them.
- × A checklist of standardised diagnostics output should be developed for assessment scientists and working groups to make reviews easier. The checklist would cover much of the output already presented in assessment

documents, such as residual and observed-expected plots. Other diagnostics, even if not included in the assessment documents, could be prepared for reviews. The following should be included where appropriate:

- observed and expected plots of survey, catch rate and size/age compositions;
 - re-runs of maximum likelihood fits from random parameter start positions to ensure that the final parameter fit is not a local maximum;
 - tests of more and less parsimonious versions of a model, providing test statistics for the exclusion/inclusion of parameters;
 - retrospective analyses, to test the predictive capability of a model.
 - parameter estimate standard errors and correlation matrix (or a cut-down version if there is a large number of parameters);
 - autocorrelations and cross-correlations of residuals for time-series models to give indications of model problems and possible improvements.
- × Assumptions and errors should be tested through simulation, where possible. Information should be presented testing the sensitivity of the results to important assumptions and errors in each assessment.
 - × Some sort of simple method needs to be developed to allow managers to assimilate risks and uncertainty in the assessments, such as decision tables. Decision tables require a definition of the decision that needs to be made and some indication of the costs resulting from the interaction between the management decision and the state of nature. Scientists and managers must collaborate in developing these tools.
 - × Recreational catches are always estimated with sampling error, and this error should be included in assessment models.

Final comments

The CIE provided me on 1 July with the almost approved versions of the panel reports (with recommendations for changes highlighted). Geographical time lag was likely the reason I did not receive it as contractually promised on 30 June. Nevertheless, I did have just sufficient time to fulfil my own contractual obligation to the CIE before my deadline of 2 July. Such a geographical time lag clearly needs to be taken into consideration in future if such tight deadlines are being set.

The main objective of the SARC-39 was to evaluate the assessments provided for the three stocks in question. I believe that the objective had already been considered by the researchers who gave the presentations, along with their immediate colleagues, and again by the working groups responsible for considering the assessments, the models applied and the assumptions made, to see if other alternatives were more appropriate or available. Nonetheless, a totally independent group such as the SARC was perfectly able to advise on the same issue, from an international perspective. Of the three stocks considered at this SARC, the assessments presented for two were considered the most appropriate at the present time, even though suggestions for further development and indeed sometimes entirely different models were made, for consideration in future. The other assessment was

rejected by the panel, although positive aspects of the work of that Working Group (for bluefish) were still identified. Overall, the efforts of all three working groups and subcommittees were exemplary and likely a reflection of the quality of assessment expertise available in the US. Hopefully, the suggestions and recommendations the panel made for future assessment will be viewed positively by the researchers. All we have done is to look at what data are available, and to advise new lines of research and analysis, including models, that could enhance the assessments in years to come, especially if the data identified as prerequisite are indeed forthcoming.

The meeting overall was conducted in excellent spirit, despite rigorous and probing debate. The panel functioned excellently as a unit, feeding off each others' strengths and abilities, and I certainly enjoyed the opportunity to talk in depth to the other three panellists, the presenters and the observers on an informal basis. I therefore wholeheartedly enjoyed the meeting and consider myself privileged to have been selected to chair it. My personal thanks are due to the CIE, who effectively organized my accommodation and facilitated the other arrangements, to Terry Smith for his efficiency in making and delivering the meeting arrangements and reference material, to Paul Rago, for supporting me with local assessment knowledge, to the other three panellists for putting up with me and my requests, and to all presenters and observers for their valuable, hugely appreciated, contributions to the meeting. Without everyone's contributions, the meeting output would not have been as comprehensive and scientifically rigorous as it turned out to be.

Andrew I.L. Payne
Chair SARC-39
2 July 2004

Appendix 1: Panelists

Chair:

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

Panel members:

Dr Michael J. Armstrong (Centre for Environment, Fisheries and Aquaculture Science, 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 (SAW 39)

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 Introduction Agenda & Conduct of meeting	John Boreman , Center Director Terry Smith , SAW Chairman Andy Payne , SARC Chairman		
Black Seabass (A) SARC Discussion	Gary Shepherd Andy Payne	Din Chen	Laurel Col
TUESDAY, 8 June (08:30 - 18:00)			
Sea Scallop (B) SARC Discussion	Dvora Hart Andy Payne	Paul Medley	Larry Jacobson
WEDNESDAY, 9 June (09:00 - 17:00)			
Bluefish (C) SARC Discussion	Andy Payne	Jessica Coakley	Mike Armstrong
THURSDAY, 10 June (09:00 - finish)			
Closed 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 Andrew Payne

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.

The Center for Independent Experts (CIE) shall provide a panel chair and three panellists for 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 seabass (*Centropristis striata*), and bluefish (*Pomatomus saltatrix*).

Specific Activities and Responsibilities

Time is of the essence for this project. The CIE's deliverables shall be provided on an expedited schedule, following the schedule of milestones in the table below. The final reports will provide key information for a presentation to be made by NOAA Fisheries at a meeting of the New England Fishery Management Council, scheduled for 13-15 July 2004. Thus, all reports must be final by 9 July 2004, to allow for time for NOAA Fisheries staff to develop their presentation.

The chair's duties shall occupy a maximum of 19 days: several days prior to the meeting for document review; the SARC meeting in Woods Hole; and several days following the meeting to review the individual panellist's Review Reports and produce the Summary Report. This report shall be a summary of the individual Review Reports, accurately and fairly representing all viewpoints. There shall be no attempt by the Chair to develop a consensus report.

Roles and responsibilities:

- (1) Prior to the meeting: review the reports produced by the Working Groups.
- (2) During the meeting: act as chairperson, where duties include control of the meeting, coordination of presentations and discussion, control of document flow and facilitation of discussion.

- (3) After the meeting: provide a Summary Report, which summarizes the findings of the individual panellist's Review Reports. The Summary Report shall be organized like the Review Reports, with an executive summary, a review of activities and, for each stock assessment reviewed, a summary of findings and recommendations that collectively emerged from the meeting. The Chair shall not attempt to reach or describe consensus on an assessment, but shall fairly summarize the individual Review Reports and draw attention to the collective conclusions and recommendations.

The Chair shall begin the summarization using the draft individual Review Reports provided by the Panellists on 21 June 2004. When these individual reports are finalized, following the CIE internal review, the CIE shall provide copies of the final versions to the Chair on 30 June for completion of the Summary Report. No later than 2 July 2004, the Chair shall submit the Summary Report to the CIE. This shall be 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 Manoj Shivlani, via e-mail to mshivlani@rsmas.miami.edu.

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

The SAW Chairman and SAW Coordinator will assist the SARC Chair prior to, during, and after the meeting in ensuring that documents are distributed in a timely fashion.

NEFSC staff and the SAW Chairman will be responsible for the production of the final SARC report, which will include the Chair's Summary Report and the individual panellist's Review Reports. Staff and the SAW Chairman will also be responsible for production and publication of the collective Working Group papers, which will serve as a SAW Assessment Report.

Contact person:

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