Fisheries and Oceans Canada / Pêches et Océans Canada
NOAA FISHERIES
NATIONAL MARINE FISHERIES SERVICE

CERT

Comité d'évaluation des ressources transfrontalières

Comptes rendus 2006/01

TRAC

Transboundary Resource Assessment Committee

Proceedings 2006/01

Proceedings of the Transboundary Resources Assessment Committee (TRAC)

Benchmark Review of Stock Assessment Models for Gulf of Maine and Georges Bank Herring

2 – 5 May 2006
Woods Hole, Massachusetts

R. O'Boyle[1], and W. Overholtz[2]
TRAC Co-Chairs

[1]Fisheries and Oceans Canada
Bedford Institute of Oceanography
Dartmouth, Nova Scotia
Canada

[2]National Marine Fisheries Service
Northeast Fisheries Science Center
Woods Hole, Mass.
USA

November 2006

Ce document est disponible sur l'Internet à : This document is available on the Internet at :
http://www.mar.dfo-mpo.gc.ca/science/TRAC/trac.html

Canada

FOREWORD

The purpose of these proceedings is to archive the activities and discussions of the meeting, including research recommendations, uncertainties, and to provide a place to formally archive official minority opinions. As such, interpretations and opinions presented in this report may be factually incorrect or mis-leading, but are included to record as faithfully as possible what transpired at the meeting. No statements are to be taken as reflecting the consensus of the meeting unless they are clearly identified as such. Moreover, additional information and further review may result in a change of decision where tentative agreement had been reached.

AVANT-PROPOS

Le présent compte rendu fait état des activités et des discussions qui ont eu lieu à la réunion, notamment en ce qui concerne les recommandations de recherche et les incertitudes; il sert aussi à consigner en bonne et due forme les opinions minoritaires officielles. Les interprétations et opinions qui y sont présentées peuvent être incorrectes sur le plan des faits ou trompeuses, mais elles sont intégrées au document pour que celui-ci reflète le plus fidèlement possible ce qui s’est dit à la réunion. Aucune déclaration ne doit être considérée comme une expression du consensus des participants, sauf s’il est clairement indiqué qu’elle l’est effectivement. En outre, des renseignements supplémentaires et un plus ample examen peuvent avoir pour effet de modifier une décision qui avait fait l'objet d'un accord préliminaire.

TABLE OF CONTENTS

ABSTRACT

The Transboundary Resources Assessment Committee (TRAC) met in Woods Hole, Mass., during 2 – 5 May 2006 to undertake a benchmark review of stock assessment models for the Gulf of Maine / Georges Bank herring stock complex. This meeting was to resolve issues with the assessment models which have been used since the last benchmark review in 2003.

RÉSUMÉ

Le Comité d'évaluation des ressources transfrontalières (CERT) s'est réuni à Woods Hole (Massachusetts) du 2 au 5 mai 2006 afin de procéder à un examen des points de référence des modèles d'évaluation des stocks du golfe du Maine. Il s'agissait notamment d'élucider les problèmes posés par les modèles d'évaluation qui ont été utilisés depuis le dernier examen des points de référence, ayant eu lieu en 2003.

INTRODUCTION

The chairs opened the meeting by welcoming the participants (Appendix 1) and providing an overview of the Transboundary Resources Assessment Committee (TRAC) process. The TRAC was established in 1998 to undertake joint US / Canada assessments of resources in the Georges Bank transboundary region. Since 2002, to enhance the level of peer review, the TRAC has endeavored to separate the review of the models and approaches used (benchmark review) from the application of the benchmark framework (assessment) to provide harvest advice. During benchmark reviews, any one of a number of components of the assessment can be considered, including the definition of the management unit, the most appropriate approach to determine stock status, reference points and projection procedures.

A benchmark review of the Gulf of Maine and Georges Bank herring complex was undertaken in February 2003, at which a number of issues were identified which required further work. These related to the determination of stock status, reference points and projection procedures. The current meeting was to consider these issues as well as undertake an assessment of the status of the complex in support of management of the 2006 fishery (see terms of reference in Appendix 2). The chairs reviewed the agenda (Appendix 3), emphasizing that it was to be used as a guide and would be open to modification as required by the discussion. Six working papers had been prepared to address the meeting terms of reference. In addition, three external reviewers (N. Cadigan, R. Mohn, and K. Piner) assisted in the peer review. These reviewers subsequently provided the chairs with their observations on the meeting and overall process (Appendix 4).

The products of the meeting were these proceedings, a Transboundary Resources Status Report summarizing the assessment and reference documents providing the technical basis of the discussion.

OVERVIEW OF ISSUES AND RECOMMENDATIONS FROM FEBRUARY 2003 MEETING

The Canadian meeting co-chair, R. O'Boyle, provided background to the meeting including an overview of the issues discussed and recommendations made at the 10 – 14 February 2003 benchmark review (Garavis, 2003). That meeting was the first joint US / Canada review of the Gulf of Maine and Georges Bank herring stock complex conducted since the International Commission for the Northwest Atlantic Fisheries (ICNAF) in 1977. It had been stimulated by a June 2000 request from the herring industries of both countries to undertake a review in TRAC. Due to data availability and competing priorities, the review could not be undertaken before the spring of 2002.

The co-chair then summarized the terms of reference of the 2003 meeting, which covered the full range of issues considered in benchmark reviews including management unit definition, resource status, reference points and projection methods.

The discussion on the management unit definition had reached a consensus that the complex should be assessed as one unit with fishing effort distributed amongst the inshore Gulf of Maine and Nantucket and offshore Georges Bank components according to estimates of the relative biomass in these areas. These estimates were to be reviewed at the current meeting. Recommendations had been made to review the US management boundaries (Figure 1), review the tagging information in the coastal New Brunswick weir fishery to refine stock affinities and initiate a study of mixing rates amongst the three components in the seasonal fisheries, all of which were underway.

Regarding resource status, the 2003 review noted reduced agreement in the aging by readers from NMFS, Maine DMR, and DFO, and had recommended that aging exchanges and workshops be undertaken to resolve these. The results of these were to be reported at the current meeting.

Few significant issues had been raised with the surveys that were to be used to monitor herring abundance trends and there had been consensus that the spring and fall NMFS bottom trawl surveys times series should be split before and after 1985.

The most important issue raised at the 2003 meeting had been the extreme divergence in the abundance estimates produced by an age-aggregated (KLAMZ) and age-structured (ADAPT) population model (Figure 2). While the ADAPT suffered severe retrospective patterns, the KLAMZ suggested the presence of older age groups to an extent not corroborated by the observations. A number of investigations had been undertaken to explain this divergence (modifications to data inputs, calculation of oldest age fishing mortality, different weights on the input components) but with no resolution. A number of recommendations had been made on post-meeting work, the results of which were to be reported at the current meeting.

An assessment of the stock complex for the 2004 fishery had also been undertaken at the 2003 meeting, using reference points calculated based on the models.

Overall, a number of significant issues had been raised at the 2003 meeting for which follow-up work had been undertaken and would be considered at the current meeting.

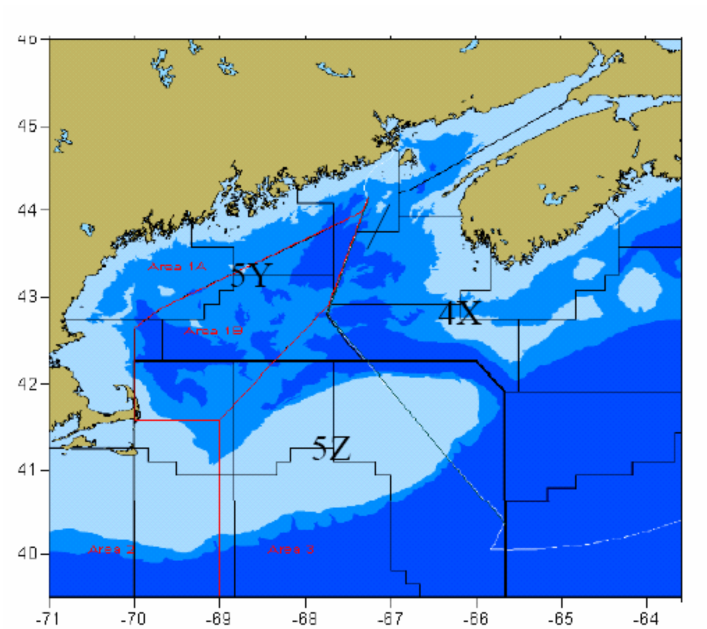

Figure 1. Management areas of the Gulf of Maine and Georges Bank herring stock complex

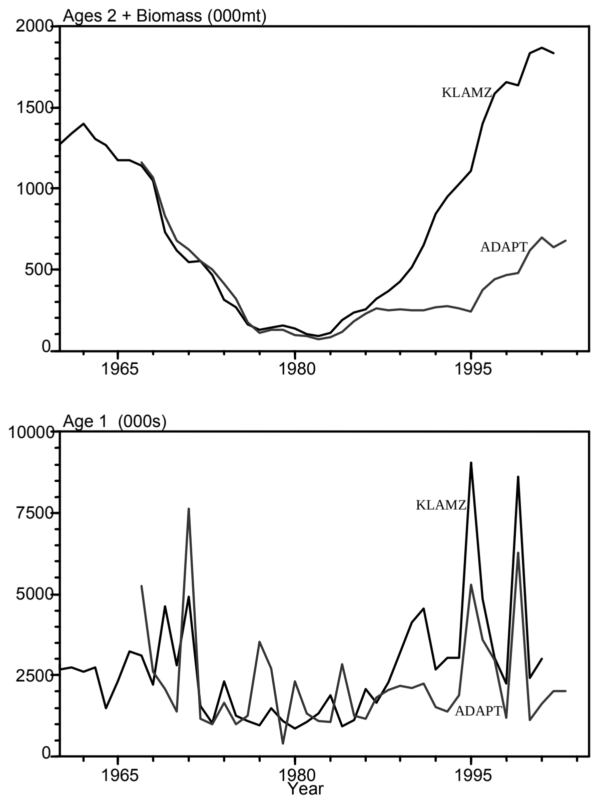

Figure 2. Age 2+ population biomass and recruitment trends produced by age-aggregated (KLAMZ) and age-structured (ADAPT) models in February 2003 TRAC meeting

DATA INPUTS

Catch at Age

Presentation Highlights (D. Libby)

A review and update of the commercial catch and sample data for Atlantic herring reported within the currently defined US waters from 1960 – 2005 was recently completed. Catch and sample data were gathered from the Maine Department of Marine Resources (MeDMR), the National Marine Fisheries Services (NMFS), the North Atlantic Fisheries Organization (NAFO) and ICNAF. Catch data were validated (verified the source of documented catch data), scrutinized and corrected for errors. The sources of data were compared and combined with regard for catch amount, period, area and gear while being mindful of duplicate data. Existing biological sample data were verified against original data sheets (MeDMR) and additional sample data were gathered from ICNAF and NAFO archives. Data were converted to common elements of date (mm/dd/yy), gear, area (lowest resolution being NMFS Statistical Areas), total length (mm), weight (gm), and age (yr). A subsequent catch at age matrix was generated from these data. It is anticipated that the updated catch and sample data will be reviewed for inclusion into future assessments and analyses.

Discussion

It was noted that the catch at age reconstruction was not in the terms of reference of the meeting. The current project is to review and reconstruct the entire catch at age, using whatever historical information exists. The catch at age may change due to the re-allocation of catches to different areas but it is likely that post-1985 changes will be minor. The new catch at age will be available in the near future; there will be a need to undertake a review of its construction.

Aging Workshop

Libby, D., J. M. Burnett, and G.D. Melvin. 2006. Proceedings of the Atlantic Herring Otolith Age Estimation Workshop, 10–11 January 2006, West Boothbay Harbor, Maine. TRAC Working Paper 2006/04.

Rapporteur: G. Melvin

Presentation Highlights (D. Libby)

An aging workshop was held at the Maine Department of Marine Resources (DMF) Fisheries Laboratory in West Boothbay Harbor, Maine during 10–11 January 2006. Attendance consisted of ten participants including four age readers.

Workshop objectives were to 1) review the current otolith processing and age reading protocols for each laboratory, 2) attempt to determine the cause(s) of the significant age estimation disparities and biases between the three laboratories and 3) develop consensus for future age estimations. The following were the Terms of Reference:

1. Assess evidence of discrepancies in Canada / US age comparisons
2. Assess herring otolith processing and age determination among laboratories
3. Assess a single method protocol in herring otolith age determination
4. Develop recommendations for future assessing of ageing among laboratories
5. Record workshop proceedings and report to 2006 Transboundary Resource Assessment Committee (TRAC)

A review of the NEFSC, DMR and SABS otolith exchange (1st Exchange) that was presented to the 2003 TRAC and the subsequent analysis was discussed. Over 200 herring otoliths were processed for age and analyzed for agreement and bias between laboratories. The minimum and maximum age estimations were two through nine. The two US laboratories were in close agreement (85%) and without significant bias as compared to SABS ages with lower agreement (75-77%) and a significant bias toward younger estimated ages beginning at 5+.

Considerable discussion focused on the otolith morphology and landmarks that were used to discern annuli. Selected otoliths from the 1st Exchange were displayed and discussed. The first discussions were on otoliths with age agreement between the three labs starting with Age 2 and progressing to Age 10. This was done to review how the three labs came to the same decision and to describe methods (which part of the otolith to read) and identify landmarks (nucleus, rostrum, anti-rostrum, outer edge, checks, winter and summer growth, etc.) in estimating age.

The group then explored the exchange sets of otoliths where there was no agreement for at least one lab. During some intense discussion, otolith artifacts such as a check between annuli one and two and how to define tightly grouped annuli near the otolith outer edge were accepted by the group.

The group decided to continue the workshop with a second otolith exchange (2nd Exchange). Approximately 200 otoliths were selected from the DMR samples that included otoliths from each month during 2004-2005. The DMR ages agreed 54% and had a significant bias toward younger ages at progressively NEFSC older ages.

The comparison of NEFSC vs. SABS showed a 39% agreement with a strong bias toward under-ageing starting at Age 4+. These results were poorer than for the 1st Exchange where there was a 76% agreement and a non-significant age bias. SABS had one reader for the 1st Exchange and that same reader read only a portion of the 2nd Exchange with another reader processing the remaining otoliths.

The comparison between SABS and DMR showed that a 58% agreement was better than compared to the NEFSC vs. DMR (54%) but considerably less than in the 1st Exchange of 78%. There was a significant over-ageing bias for the DMR reader for 5+ ages.

The poorer results of the 2nd Exchange compared to the 1st were a surprise and somewhat disheartening for the group. The otolith comparisons and discussions during the workshop were positive and everyone felt optimistic about the 2nd Exchange resulting in better agreement and less bias compared to the 1st Exchange. Reasons for the poor results could have been 1) change in readers and possible ageing continuity loss at the NEFSC and SABS laboratories, 2) unfamiliarity of trying to read otoliths that another laboratory has processed and mounted, and 3) being presented with fish of different populations, growth rates than the usual fish that the particular laboratory processes.

Recommendations for follow-up activities included:

- Continuation of collaboration amongst laboratories to develop like age estimations using herring otoliths
- More and regular otolith exchanges and analysis
- Acquisition of useful and efficient validation methods
- Establishment of a permanent and dynamic herring otolith reference collection developed by all ageing partners

Discussion

It was noted that the otoliths used in the 1st 2003 exchange were not reread in the 2nd exchange. It was suggested that this be done to confirm the observed inter-laboratory differences.

It was asked if the length of a herring was available to the ager at the time of the aging. Yes but there has been a large compression of lengths in recent years as evidenced by the growth rings being closer together.

It was confirmed that the otoliths from the 1st and 2nd exchange came from the same areas. It was emphasized that the results of the 2nd exchange are still preliminary and have not benefited from inter-ager dialogue on the readings, which had occurred in the 1st exchange. Overall, the reasons for the inter-laboratory differences are still to be determined.

There then followed a discussion on the potential impact that the aging workshop findings would have on the assessment. The DFO aging is used to construct the Canadian weir catch at age, which currently represents about 10% of the total catch (10,000 t) with 10% of this being older fish. It was about 50% of the total catch historically but even then, the weir catch was composed mostly of smaller fish, which are easier to age. The DFO aging is also used for the DFO February bottom trawl survey. DMR has provided the aging for most of the catch at age, while NMFS has provided the aging for the NMFS surveys. Regarding consistency over time, the DMR aging is the most consistent. This was the case for the NMFS aging prior to 2005 but in that year, the ager changed, thus creating the possibility of an inconsistency, which needs to be investigated.

Overall, from examination of the catch at age, there is no evidence, other than the 2nd aging exchange, of inconsistency of aging through to 2004. If aging is inconsistent, it is likely to be occurring in the older age groups, certainly beyond Age 4 and potentially Age 6. If the aging information were to be used in the assessment, consideration would need to be given to aggregating the information on the older ages e.g. create an Age 6+ group.

Landings Numbers and Weights at Age

Overholtz, W.J., L.D. Jacobson, M. Cieri, and C. Legault. 2006. Assessment of the Gulf of Maine – Georges Bank Herring Complex, 2005. TRAC Working Paper 2006/02.

Rapporteur: J. Grist

Presentation Highlights (W. Overholtz)

The 2002 and subsequent landings number and weight at age in the 2003 assessment (Overholtz et. al., 2004) were preliminary and thus were updated. The process on how this was done was briefly described.

Discussion

It was requested that Table 3.2 of the 2003 assessment be produced during the meeting. This provides the landings by stock component and year. This is included here as Appendix 5. The need for a graph of total landings over time was also noted. This was produced later during the meeting as well.

Groundfish Surveys

Overholtz, W.J., L.D. Jacobson, M. Cieri, and C. Legault. 2006. Assessment of the Gulf of Maine – Georges Bank Herring Complex, 2005. TRAC Working Paper 2006/02.

Rapporteur: J. Grist

Presentation Highlights (W. Overholtz)

The sampling coverage from the Mid-Atlantic to Georges Bank and the Gulf of Maine (GOM) of the three NMFS stratified random bottom trawl surveys (winter, spring and fall) was described. The winter survey does not cover the Gulf but does sample Georges Bank. The herring abundance trends were also described. The rise in abundance in recent years is consistent with the expansion of the age range into the older age groups (7-10) observed in all three surveys. It was noted that the decline in the abundance of the middle ages during the 1980s – 90s was more dramatic in the spring as opposed to the fall survey.

There then followed a description of the spatial distributional changes in the resource coincident with abundance, emphasizing that the offshore component was severely depleted in the 1970s with the recovery first occurring in the inshore, and progressing west to east. By 1998, almost all the GOM was occupied by herring, both inshore and offshore.

In contrast to the trawl surveys, the acoustic surveys for these years indicated that there was a progressive increase in the biomass index through 2001, with a major decline in 2002, although there is no evidence of this in the catches of the fall survey. In 2003, the acoustic survey index declined further, while the fall bottom trawl survey observed major groups of herring outside of the acoustic survey area. The same occurred in 2004 - acoustic survey index down while those of the bottom trawl surveys showed an expanding distribution. In 2005, again there was a low estimate of abundance from the acoustic survey but herring were distributed throughout the GOM proper.

Discussion

It was asked if during the early 2000s, herring were more abundant in the inshore versus the offshore. The presentation was a perspective of the whole complex over the last decade based on the bottom trawl surveys, to contrast to the acoustics survey to be considered in the next presentation.

Is there any information on differences in age composition in the fall by spawning component? This has not been examined. It would be necessary to consider the age composition in the spawning aggregations in inshore GOM, Nantucket and on Georges Bank along the 100m contour. It was pointed out that the fall bottom trawl survey is sampling throughout the stock complex's distribution during late September – mid October, coincident with the timing of the acoustic survey.

There followed an extended discussion on the seasonal movements and spawning locations of herring, which will not be recounted here.

The concern was raised that the winter survey might not be sampling the whole stock. However, the inshore and offshore herring move south in the winter and thus the survey is sampling both components. It was then pointed out that the DFO February bottom trawl survey would be operating too far north to representatively sample the herring complex. This is a consideration when discussing model input data.

It was asked what age groups are available during the survey periods. During the spring NMFS survey, Age 2+ would be available while during the fall NMFS survey, Age 3+ would be available. It was pointed out that during the fall spawning period, the herring are close to the bottom, which increases the availability to the trawl gear. Variability in the survey results was then discussed. Coefficients of Variation (CVs) are generally in the order of 20-40%. Notwithstanding the issues with the bottom trawl surveys, it was pointed out that the expansion of range of the herring complex over the Gulf of Maine is consistent with the abundance trends from the surveys and increases in the proportion of older herring.

There was some discussion on catch rate (CPUE) information from the fishery. There have been a number of changes in gear, vessels, and so on. As well, CPUE for herring generally suffers from hyperstability (catch rates not proportional to biomass). For these reasons, commercial CPUE is not considered a reliable index of abundance.

It was noted that it would be useful to undertake smoothing of the survey information to summarize the trends. This was undertaken later in the meeting and is reported below (Figure 3).

It was requested that the CVs be included in the data summaries as a matter of practice.

Acoustic Survey

Overholtz, W.J., L.D. Jacobson, M. Cieri, and C. Legault. 2006. Assessment of the Gulf of Maine – Georges Bank Herring Complex, 2005. TRAC Working Paper 2006/02.

Rapporteur: J. Grist

Presentation Highlights (W. Overholtz)

The offshore has been surveyed acoustically using the NMFS survey vessel Delaware II since 1998, generally at the same time and covering the same area where spawning herring are occurring as the fall survey. The survey series is consistent during 1999 – 2005 which show a consistent decline in herring abundance since 2001, with no identifiable mechanism to explain this decline. However, during the first three years of the survey, the majority of herring sampled were in a developing stage and were cueing to spawn. There was no sense that herring were moving into the area. This changed in 2002 with a larger proportion of the herring active and in the resting or ripe stage. Overall, in the first three years of the survey, there was little herring activity. This is completely different from the last four years of the survey when the fish were active and in different spawning condition. Based upon these availability changes, the acoustic survey is not considered to provide a reliable indicator of abundance.

Discussion

There was a discussion on the timing of the acoustic survey in relation to spawning. A week or two can make a difference. Yes, the survey is conducted at the same time but if the herring behavior changes (e.g. temperature influence – the last three winters were cold), there could be an effect. There was then a discussion on the survey coverage in relation to the overall herring distribution. There was a sense that the northern part of the population was moving in and out of the survey area in the last four years, violating the assumption that the population was being fully surveyed. Maps of the distribution of herring during the NMFS fall survey might illuminate this issue but the movement likely occurs too fast.

There was further discussion on possible processes to explain the decline in the acoustic survey since 2001 (fish behaviour, changing maturity schedule, distribution, etc) but there was no consensus. Thus, while the trend appeared inconsistent with other sources of information, there was no compelling evidence to exclude the acoustic survey information.

Independent Information

Overholtz, W.J., L.D. Jacobson, M. Cieri, and C. Legault. 2006. Assessment of the Gulf of Maine – Georges Bank Herring Complex, 2005. TRAC Working Paper 2006/02.

Rapporteur: J. Grist

Presentation Highlights (W. Overholtz)

A number of datasets corroborate the abundance trends observed in the bottom trawl surveys. The average size at age of a herring in the fishery has declined since 1985. While the trends at Age 2 might be due to the fishery, those for the older ages may be density dependent. The percent herring in the diet of predators (dogfish and cod) has increased since 1985 with perhaps some leveling off in the recent period. Before then, sand lance was more common in the diet. The consumption of herring by marine mammals, large pelagics, sea birds, fish and so on has also increased, independent of the assessment, suggesting that there is a substantial abundance of herring supporting these predators. Further, in both the spring and fall surveys, there has been a clear increase in the proportion of sets that have caught herring. All these indicators suggest that herring in the complex are bountiful.

Discussion

There was discussion on size at age differences by stock component. It appeared that there have been steeper declines on Georges Bank which might suggest that density dependent processes were occurring. However, this could also be due to fishery effects, particularly at the younger age groups. How the sampling was undertaken and analyzed was then discussed. While the fishery could be an issue, the trends were felt likely to be biologically based. There was discussion on how high the biomass would have to be to depress growth rates by 20% or so. It was agreed to revisit this during the model discussion.

Overall, there was consensus that herring abundance has increased since the mid-1980s. The issue is by how much compared to the 1960s. It was noted that herring is a key prey for a number of species, including dogfish. There has been an increase in herring in the diet of dogfish since the late 1980s, corroborating an increase in herring abundance. This led to an extended discussion on recent consumption estimates of herring by marine mammals, large pelagics and seabirds. These estimates have generally tracked herring abundance and currently suggest substantial amounts of herring. The consequences for natural mortality (M) and its relationship with fishing mortality (F/Z ratio) were discussed. Examining this ratio might provide information on herring abundance changes. Different methods to consider this (e.g. landings / F ratio) were discussed with no consensus. It was agreed to revisit this during the model discussions.

The last piece of independent information presented was the incidence of herring in the groundfish surveys. There was a clear increase in both the spring and fall surveys, matching the trend in the kg per tow.

Overall, the independent information suggests that the herring resource has increased since the mid-1980s.

Acoustic Survey Disaggregated by Age Index

Melvin, G.D., and W. Overholtz. 2006. Gulf of Maine – Georges Bank Acoustic Biomass Index - Age Disaggregated. TRAC Working Paper 2006/03.

Rapporteur: J. Grist

Presentation Highlights (G. Melvin)

In the 2003 assessment, the acoustic survey time series was short (1999 – 2002) and could only be used to gauge absolute biomass. With three more years of surveys and a consistent parallel transect survey design and protocol, it is now possible to develop an age – disaggregated relative index. The procedure by which this is attained was described, emphasizing that the backscatter is converted to numbers per size interval which is then applied to an age/length key (ALK) to produce numbers at age. An examination of the resultant survey abundance at age suggested that the survey is tracking year-classes.

Discussion

It was asked if the analysis used annual age/length keys to which the answer was yes, the ALKs from the NMFS surveys were used. These are applied to length frequency sample information collected by mid-water trawl during the acoustic survey. It was then observed that the acoustic survey catches relatively large, Age 7 to Age 9 herring compared to the trawl survey. This is expected as the acoustic survey covers the area where the larger spawners are found. It was clarified that the bottom trawl survey ALK is based on fork length which is converted to total length for use with the acoustic survey.

The age changes observed (expansion of age range) do not explain the apparent decline in the acoustic survey abundance index since 2001. If a true decline were occurring, age span contraction would be expected, not the reverse. There have been some changes in maturation that suggest that the availability of spawners to the survey might be an issue but there is no corroborative information on this.

In summary, there was general agreement that the herring resource has increased from the 1980s to the early 2000s. It is the trend since then that is at issue. While the acoustic survey suggests a decline, this might be due to changes in survey catchability. However, none of the surveys show large increases – just stability or a modest decline. It was agreed to include the acoustic disaggregated index up to Age 6+ in the model to evaluate its contribution to the assessment.

INITIAL MODEL INVESTIGATIONS

Virtual Population Analysis (VPA)

Overholtz, W.J., L.D. Jacobson, M. Cieri, and C. Legault. 2006. Assessment of the Gulf of Maine – Georges Bank Herring Complex, 2005. TRAC Working Paper 2006/02.

Rapporteur: M. Power

Presentation Highlights (W. Overholtz)

Initial investigations using VPA were attempted using ages two to 10+ with ages three to nine estimated and ages seven to nine assumed to be fully recruited. Three models were conducted. The first included all the surveys (winter, fall, spring, DFO & acoustic) and showed a decline in abundance during 2001 – 05, this thought to be driven by the acoustic survey. The acoustic survey was dropped from the second model, which also exhibited a decline in abundance during 2001 – 05 but this was not pronounced. The last model did not include both the acoustic and the DFO survey and showed an increase in spawning biomass since the 1980s. All runs exhibited very high fishing mortalities (2.0) in recent years.

Discussion

It was confirmed that the surveys all received equal weighting.

Age Structured Assessment Program (ASAP)

Overholtz, W.J., L.D. Jacobson, M. Cieri, and C. Legault. 2006. Assessment of the Gulf of Maine – Georges Bank Herring Complex, 2005. TRAC Working Paper 2006/02.

Rapporteur: M. Power

Presentation Highlights (W. Overholtz)

The presentation started with a brief explanation of the Age Structured Assessment Program (ASAP, Legault and Restrepo 1999) The ASAP model uses a forward solving approach to estimate population abundance and fishing mortality rates from catch at age and tuning index data. The model was formulated for Atlantic herring using all available information for 1967 to 2005 and ages 2 to 6+ with the exception of the Canadian research vessel (RV) survey and the larval surveys. The stock-recruitment relationship was not assumed and recruitment estimated as a mean with deviations about this. The numbers at age in the first year, 1967, were freely estimated. Changing the weights assigned to different parts of the objective functions allows the model to be more VPA-like (emphasizing catch at age information) or more KLAMZ-like (de-emphasizing catch at age information). The catch at age data were treated as multinomial distributions and input effective sample sizes compared to resulting effective sample sizes. The input effective sample size for all years was varied from 10 (KLAMZ-like) through 50 (Base Case) to 200 (VPA-like) while holding all other objective function weights the same. The tuning indices were all equally weighted and input as age-specific series. Error distributions for the indices were assumed lognormal. Fishery selectivity was set to one for ages 2+. Catch weights at age were used to match the total catch in weight with high precision, while the match to the catch at age and tuning indices was determined by the emphasis placed on each. Natural mortality was set to 0.2 for all ages. Maturity at age for all years was set at 0.21, 0.86, 0.93, 0.98, and 1.00 for ages 2 through 6+ respectively. Spawning stock biomass was calculated as the sum of numbers at age times maturity at age times weight at age. Total biomass was calculated outside the model using weights at age for the start of the year times numbers at age at the start of the year. The Hessian matrix was used to estimate uncertainty in the spawning stock biomass time series. Time constraints did not allow the uncertainty estimation using the Markov Chain Monte Carlo (MCMC) procedure, although previous experience has shown that the Hessian estimates are often quite good. Retrospective analyses were conducted by sequentially removing the terminal year of data and re-estimating all parameters. The ASAP model is available at the NFT website (http://nft.nefsc.noaa.gov/).

There then followed a description of the three models attempted. The first incorporated all survey information and showed a decline in the resource since 2001. Large differences in the 1967 abundance estimate with the VPA were observed. When a stock / recruitment relationship was fit, consistent with theory, that of Beverton and Holt appeared useful. In the second model, the acoustic survey was not used but none of the fits changed a great deal. There was a decline in abundance since 2001, again with a large difference with the VPA in 1967 and a large retrospective pattern. The last model did not include the acoustic and DFO survey and showed increasing spawning biomass after 2001 and a small retrospective pattern.

Discussion

There were a few questions of clarification on the details of the model inputs.

KLAMZ

Overholtz, W.J., L.D. Jacobson, M. Cieri, and C. Legault. 2006. Assessment of the Gulf of Maine – Georges Bank Herring Complex, 2005. TRAC Working Paper 2006/02.

Rapporteur: M. Power

Presentation Highlights (W. Overholtz)

The KLAMZ model as used in the 2003 assessment (Overholtz et. al, 2004) was updated, using the same structure as in 2003 but with some modifications. Ages 2 and 3+ were modeled using as calibration indices (internally weighted) the following surveys: NMFS winter, spring and fall, DFO February, commercial acoustic, NMFS larval (up to 1994) and Canadian larval (up to 1995). The catchability of the acoustic survey was entered as a prior of 1.0. The catch series and thus modeled population processes started in 1959 which is earlier than used in the age structured models. The results, including a Fox surplus production model (preferred) were presented. It showed high recent biomass (1.8 mt in 2005) with surplus production greater than landings since the early 1990s. There was a slight retrospective pattern in this initial model. A second model was tried without the acoustic survey and provided similar results. A third model without both the acoustic and DFO surveys again produced similar results. A final formulation was presented with the cacthability of the acoustic survey estimated and not set at 1.0. This produced high recent fishing mortalities due to the model fitting the current declining population trends in the acoustic survey.

Discussion

There was clarification of a number of the constants and parameters used in the models. The acoustic survey catchabilities illustrated some interesting patterns that were discussed. This led to a discussion on the differences between the models. VPA is strongly influenced by the catch at age (CAA) with the stock/ recruitment relationship an output rather than an input. ASAP on the other hand does not try as hard to fit the CAA with more weight given to fitting the survey trends. KLAMZ does not use the age data, relaying completely on the age aggregated survey trends.

VPA Explorations

Gavaris, S. 2006. Exploration of VPA Model Formulations for Gulf of Maine / Georges Bank Herring Complex. TRAC Working Paper 2006/05.

Rapporteur: M. Power

Presentation Highlights (S. Gavaris)

Results from two assessment models gave divergent results at the 2003 assessment review. However the basis for the differences was not fully explored. An age structured framework using alternative model formulations that gave results spanning the range of those obtained at the 2003 assessment was employed for these investigations because it facilitated comparison of assumptions and diagnostics. Some features of the data (e.g. abrupt decline at older ages in the early 1970s, high catches of Age 2 during the mid-1970s to mid-1980s followed by negligible catches from these year classes subsequently (apparent change in survey catchability around 1985) present major challenges in fitting age structured models to these data.

Initial exploration using catch and indices out to Age 10 indicated that the change in survey catchability in 1985 coincided with the beginning of recovery following a decade of collapse. Accordingly, these weak or nearly absent year classes do not track well before and after 1985, resulting in very weak linkage of model results pre and post-1985. This precluded establishing the scale of the post-1985 population increase to the population magnitude observed pre-1985. While the pre-1985 population decline and the population magnitude during the collapse in the late 1970s to mid-1980s was relatively robust to model assumptions, the scale of the post-1985 population increase was very sensitive to model assumptions. Population biomass in 2006 could vary by a factor of about five, depending on the model assumptions.

Subsequent explorations using catch and indices out to a 6+ age group were used to investigate if some discrepancies and contradictory signals could be resolved. While aggregating to a 6+ age group improved some diagnostics, the main issue in the assessment, estimating the scale of the post 1985 recovery, remained problematic.

It was concluded that the information content of the data regarding the scale of the post-1985 population increase was very low. Informed expert judgment on subtle features of the diagnostics (e.g. survey catchability patterns by age, fishery partial recruitment patterns by age, comparison of proportional survey increase from late 1980s to 2006 versus proportional population increase) would need to be invoked to select suitable model assumptions.

Discussion

There were questions on the veracity of the CAA at the older ages. This led to discussion on the calculation of the oldest non-plus age group fishing mortality (11). It was noted that the survey catchabilities up to Age 6 appeared to be reasonably well behaved but not older. For the older age groups, the CAA did not seem to be informative. Small changes produced large differences in output. The models would benefit from external information on the expected trends in partial recruitment and survey catchability. For instance, what has been the trend in fishing effort since 1980? Total fishing effort has increased from 1987 onwards with the fleet growing from one to several large trawlers. On the other hand, the inshore stop seine fishery no longer exists. While there was no information available on effort trends and recognition that management has considerably affected fishing activities, it was felt that overall effort was more focused on the inshore area with little on the offshore.

Retrospective Investigations of VPA

Cadigan, N. 2006. Local Influence Diagnostics for the Retrospective Problem in the Gulf of Maine – Georges Bank Herring VPA of 2005. TRAC Working Paper 2006/06.

Rapporteur: M. Power

Presentation highlights (N. Cadigan)

A substantial retrospective problem exists in some of the VPA formulations published (Overholtz et al., 2004) for the Gulf of Maine - Georges Bank Herring complex (GMGBH). We applied local influence diagnostics to assess possible causes of the retrospective patterns. The VPA formulation we investigated is very similar to the final run in Overholtz et al. (2004). The retrospective diagnostics are described in Cadigan and Farrell (2005). They can be used to find perturbations to model inputs that remove or greatly reduce retrospective patterns. The magnitude of the perturbations required to remove retrospective patterns can be used to assess the plausibility that the input is the source of the problem. We consider four distinct perturbation schemes on inputs of fishery catches, natural mortality, survey catchability, and extrinsic estimation weights (case weights) in order to determine if the retrospective patterns in the GMGBH VPA are more likely caused by any of these components.

Results indicated that none of the perturbation schemes were able to completely remove retrospective patterns although it was possible to reduce ρ, a common retrospective metric described by Mohn (1999), to zero. Age-specific patterns were usually not improved. Large perturbations were required to reduce ρ to near zero. The VPA fit was poor at older ages, perhaps because of aging errors, and a better model formulation involving a plus group should be considered. A diagnostic analysis of such a formulation may be more revealing.

Discussion

It was pointed out that lack of a retrospective pattern does not necessarily mean that the model is more accurate. One can have a model with little or no retrospective pattern but is a poor fit to the data. A retrospective pattern is evidence of a poor model but when is not present, it is still necessary to consider the residual patterns and other model diagnostics.

FURTHER MODEL INVESTIGATIONS

Rapporteur: M. Power (all), R. Mohn (VPA), K. Piner (ASAP), and N. Cadigan (KLAMZ)

Based upon the presentations of the initial model formulations, there was a general discussion on the data inputs and the model structure with the overall aim being the production of one assessment model for use as the basis of management advice. These discussions were used to guide subsequent model explorations conducted at the workshop.

Data Inputs

To facilitate comparison amongst the model outputs, it was considered important to define a consistent set of data inputs. The decisions on these are provided below.

Catch at Age

Due to both aging uncertainties and sampling issues, it was agreed that age structured formulations (VPA and ASAP) should use ages one to five with a six plus group. Concerns were raised on the patterns in the weights at age, but these require investigation after the workshop.

Survey Indices

It was agreed to drop the DFO bottom trawl survey from the formulations as this survey does not cover the entire stock area.

The winter, spring and fall NMFS surveys were to be included in the models as was the acoustic survey, all as relative indices of abundance. It was recognized that there are issues with the acoustic survey series but these were not considered sufficient to exclude it from the models.

It was agreed to include the Canadian and US larval survey time series.

Model Structure

There was agreement of a number of aspects of the model formulations with suggestions on further explorations.

Population Processes

It was agreed that the formulations would be investigated without imposing a stock / recruitment relationship.

Regarding natural mortality (M), 0.2 would be used in the formulations with likelihood profiling used to investigate alternative values. As well, it was suggested to compare the landings/total removals (based upon the consumption information) ratio with the F/Z ratio from the models.

Survey Processes

It was agreed to estimate survey catchability (Q) by age as well as split the time series into pre and post-1985. It was suggested that the VPA formulations investigate the impact of dropping the pre-1985 data.

Fishery Processes

Regarding the partial recruitment, it was agreed that it would be assumed to be flat topped after a certain year, dependent upon examination of the fishing mortality (F) matrix. Estimation of the F on the oldest age would be investigated for blocks of years.

Fit

The relative weight given to the CAA could be examined in ASAP through a sensitivity analysis. Regarding the surveys, it was suggested to conduct model optimizations without weighting and consider the mean square residuals (MSR) before employing self weighting (intrinsic).

Comparison of Models and Plausible Outputs

There was discussion on the need to determine how best to compare the model outputs. A number of products were identified. It was agreed that F / B profiles for 2005 and the last six years would be useful in portraying the overall results of the models. It was also considered useful to plot model-derived surplus production as a function of biomass. It was noted that ASAP provides a predicted age structure for 2005 which could be compared to that observed. Where possible, likelihood profiles of selected parameters (e.g. M) should also be investigated. The issue of plausibility of the various model outputs was discussed given the information that was available. What aspects of the data need to come out in the results? These split into two categories: biomass and fishing mortality related issues.

Regarding biomass, it was agreed that the scale of increase in the Age 2+ biomass from 1985 to the present as indicated by the NMFS winter, spring and fall surveys was important to reflect in the model output. Another feature was the low Age 7+ numbers during 1980 – 1990. Smoothing of the three NMFS Age 2+ kg/tow time series (Figure 3) suggested a 2005/1985 ratio in the order of 2.5 – 3.5 with an average of 2.6 with stability since about 1996.

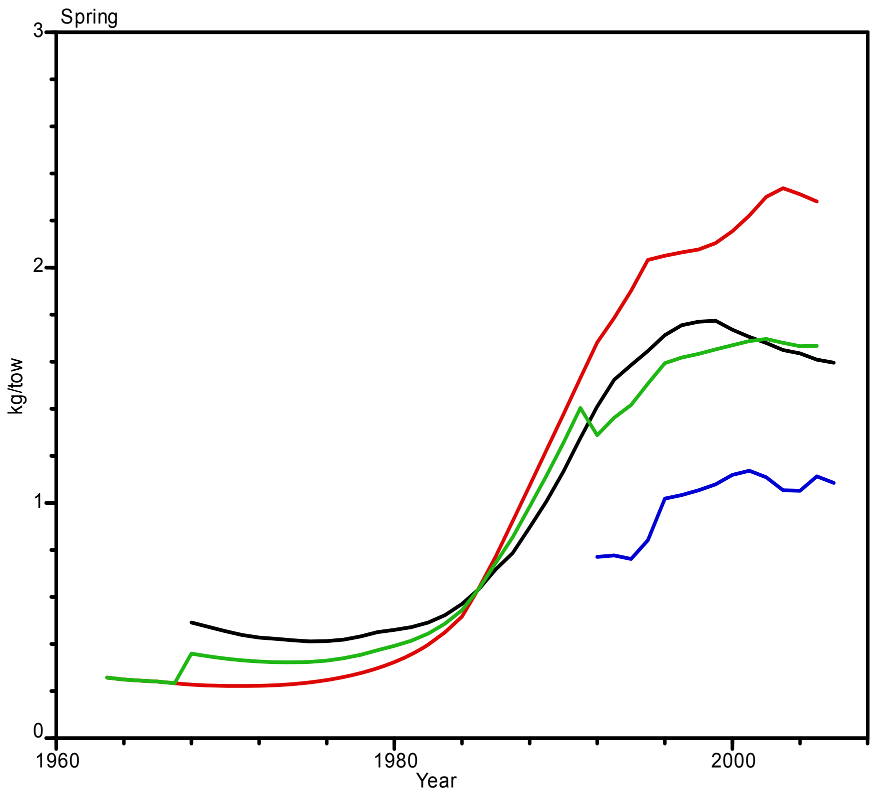

Figure 3. NMFS surveys (red – fall, black – spring and blue – winter) smoothed and normalized to their means. These were combined into a single series (green line) by averaging across the three series. The smoothing was done with a kernel smoother with a weight subjectively set to retain some of the variation in the series. The ratio of 2005 to 1985 for the combined estimate is 2.6.

Regarding fishing mortality, the high fishing mortalities observed in the 1970s should be reflected in low Age 6+ numbers during 1980-90. It was agreed that a flat topped partial recruitment was more plausible than a dome shaped pattern for a herring fishery. If a flat top was imposed in the ASAP, recent observed and predicted CAA could be examined for anomalies. There were a number of suggestions to consider Catch / Biomass ratios from the surveys both pre and post-1985, along with fishing effort trends and age composition to determine if current F is consistent with these trends in relative fishing mortality.

Regarding model fit, it was agreed that the presence of a retrospective pattern in both spawning biomass and fishing mortality was considered undesirable as was high mean square residual (MSR).

These observations were considered in explorations of the various models, a summary of which is reported in Table 1 below.

ASAP

Four ASAP models were presented. Model 1 constituted the Base model. It did not fit a stock/recruitment relationship and set partial recruitment (PR) for ages 2+ at 1.0 (flat topped). The age – specific survey indices received equal weighting. The catch at age received a weighting of 50 in this model. Model 2 was a sensitivity analysis to investigate what would occur if the fishery selectivity was estimated. It used the same structure as Model 1 except that one partial recruitment pattern was estimated as the entire time period. Model 3 used the same structure as Model 1 but was VPA-like with respect to its emphasis on the age composition of the catch. Here the weighting was 200 instead of 50. Model 4 used the same structure as Model 1 but was KLAMZ-like with respect to its de-emphasis of the age composition of the catch (weighting of 10) and thus relatively more weight being given to the survey indices.

A comparison of the results of the Base, VPA-like and KLAMZ-like models with the plausible outputs discussed above is provided in Table1. The results of the PR sensitivity analysis are not provided as this model produced a partial recruitment which declined with age and was not considered realistic.

All three models had 1967 Age 2+ biomass in the order of 1600-1700 kt declining to about 230-280 kt in 1985 and rising thereafter. All showed trends similar to those in the surveys since 1996. What discriminated them was the extent of the post-1985 trend, with the KLAMZ – like model estimating a 5.6 times increase in Age 2+ biomass during 1985 – 2005 and the other two models 3.3 – 4.4 times increases.

Regarding fishing mortality, all three models estimated low Age 6+ numbers during 1980-1990 as expected. There were differences in the comparison of the 2002 – 2005 observed and predicted catch at age. All three showed patterns in the residuals for Age 2 – 5 and particularly Age 6+ with predictions of too many old fish. The worst patterns were in the KLAMZ – like model and the closest match in the VPA – like model. The recent fishing mortality patterns also differed with the Base and KLAMZ – like models providing a step down pattern during 1985 – 2005 while that in the VPA – like model was more gradual. The 2005 fishing mortalities ranged from 0.08 for the KLAMZ – like model to 0.17 for the VPA – like model.

Regarding model fit, for fishing mortality, while the retrospective pattern was still present, it was not as apparent as that for spawning biomass. The most significant pattern was observed in the VPA – like model and the least in the KLAMZ – like model.

Overall, the VPA and KLAMZ – like models illustrated the tradeoff in the weighting placed on the catch at age and the survey data. Putting more weight on the catch at age results in both a lower recent and smaller post-1985 increase in Age 2 + biomass. De-emphasizing the catch at age results in both a higher recent and greater post–1985 increase in Age 2+ biomass. Within the same modeling framework, the difference in results between the KLAMZ (age-aggregated) and VPA (age-disaggregated) models are illustrated.

During the discussion of these models, it was suggested that a likelihood profile of natural mortality be conducted using the Base model. It showed that M in the order of 0.3 – 0.4 provided better fits, suggesting that the assumed value of 0.2 might be too low.

Two Surplus Production Models (Schaefer and Fox) were produced from the Base model. Bmsy was 858 kt and 783 kt for these two models respectively, while the companion MSY estimates were 180 kt and 161 kt. Fmsy was 0.21 and 0.20 for the two models, which is very close to the assumed value of natural mortality and considered plausible.

VPA

The earlier VPA model was reformulated to include ages 1–5 and 6+. Fishing mortality on Age 5 was set to that on Age 6+ for 1975 to present. Before then, one Age 5/6+ F ratio was estimated (at about 0.5). A flat-topped partial recruitment was assumed.

The Age 2+ biomass from this model declined from 1200 kt in 1967 to 346 kt in 1985 and thereafter rose to 519 kt in 2005. This represented a 1.4 times increase in biomass during this period, compared to the 2.5 – 3.5 increase deemed plausible. The post-1996 pattern was not as stable as that observed in the ASAP models.

Regarding fishing mortality, this model also predicted low Age 6+ numbers during 1980–1990, but a higher 2005 F compared to those produced by the ASAP models. Compared to the latter, the post–1985 trend in F was relatively stable around 0.25.

The MSR was 1.7 compared to the earlier VPA model (0.7) which is not unexpected given the imposed change in age range. However, the survey indices were not well fit with residual patterns apparent. For the spring and fall NMFS surveys, the model first under-predicted and then over-predicted Age 6+ numbers in recent years. Age – specific patterns in the survey catchabilities were also seen with those of the acoustic survey showing a monotonic increase with age. The retrospective pattern for biomass appeared to be more severe than seen in the ASAP models with speculation that this might be due to the influence of the acoustic survey.

No Surplus Production Model was attempted using the results of this model.

KLAMZ

The KLAMZ model was attempted with the stock / recruitment both estimated and not estimated. The latter formulation had significant problems fitting the data and produced very high recent biomass levels. This was a useful comparison as it highlighted the need for information on year class strengths, either obtained through use of aging data (ASAP and VPA models) or imposition of a stock/recruit relationship (KLAMZ).

The trawl door effect of the spring and fall surveys was higher on Age 2 than Age 3+ with the catchabilities showing the reverse trend. The estimated catchability of the acoustic survey was 0.71, showing that this survey is not an absolute indicator of abundance.

Age 2+ biomass from the KLAMZ model without the spring Age 3+ covariate for the trawl doors was about 1300 kt in 1967, declining to 258 kt in 1985 and then increasing to about 1000 kt (an increase of 3.8 times) in 2005. Recent trends in fishing mortality from the model indicated a decline since 1985 to the 2005 value of 0.10. There was a comparison of the trends in the removal/consumption and F/M ratios which appeared to be similar. The trend in the catch/spring and fall survey biomass ratios were not considered informative due to the size composition differences between the two data sets.

The fit of the Age 3+ to the spring survey series was poor in recent years with large retrospective patterns seen in both the beginning (1959) and end (2005) of the time series.

A Schaefer Production Model using this KLAMZ model provided a BMSY of 1100 kt, an MSY of 150 kt and and Fmsy of 0.15.

Table 1. Comparison of model results with agreed - to plausible outputs

Component	Performance Measure	ASAP Base	ASAP VPA-Like	ASAP KLAMZ-Like	VPA	KLAMZ without spr 3+ covariate
Age 2+ Biomass	2.5 – 3.5 times increase in RVs during 1985 – 2005	4.4	3.3	5.6	1.4	3.8
	Stable trend since 1996	Similar to RVs	Similar to RVs	Similar to RVs	Bit of up & down	Not considered
	1967 / 1985 / 2005 kt trend	1600 / 230 / 1040	1720 / 230 / 760	1590 / 280 / 1560	1200 / 346 / 519	1319 / 258 / 987
Age Structure	Low Age 6+ numbers during 1980 – 1990	Yes	Yes	Yes	Yes	NA
	Predicted & Observed CAA in recent period; Age 2 – 5	Problems apparent	Not considered	Not considered	NA	NA
	Predicted & Observed CAA in recent period; Age 6+	Predicted 6+ higher in 2002 -2005 bad	Predicted 6+ higher in 2002 -2005 Closest match	Predicted 6+ higher in 2002 -2005 worst	NA	NA
Fully Recruited Fishing Mortality	2005 F	0.10	0.17	0.08	0.25	0.10
	Trend since 1985	0.3 step down in 1990 to 0.1	0.3 more grad to 0.17	0.2 step down in 1990 to 0.08	Stable at around 0.25	decline
Fit	MSR / Residuals				MSR 1.7 considerably higher than earlier model Residuals Spring & fall RV 6+ neg (under pred) to pos (over pred) Monotonic increase in age – specific Acoustic survey Q	Poor 3+ fit at end of spring series Less wt on S/R, causes model to fit fall 2 with little change in other fits
	Retrospective SSB	Large	Larger than in Base	Small overall and Little in 2004/05	Large retro; looks more severe than ASAP	Large retro (smaller in 1959 & larger in 2005) at both ends
	Retrospective F	Present but not as apparent as for SSB	Present but not as apparent as for SSB	Present but not as apparent as for SSB	Present but not as apparent as for SSB	Present but not as apparent as for SSB

Comparison of Biomass / Fishing Mortality Profiles from Terminal Year of Models Investigated

It was considered informative to display the results of all the models together to gain an appreciation of how each is portraying resource status (Figure 4). In general, the VPA and ASAP VPA – like results were comparable while those for KLAMZ and ASAP KLAMZ – like were as well. Overall, the ASAP base model fit between the extremes of the other two approaches

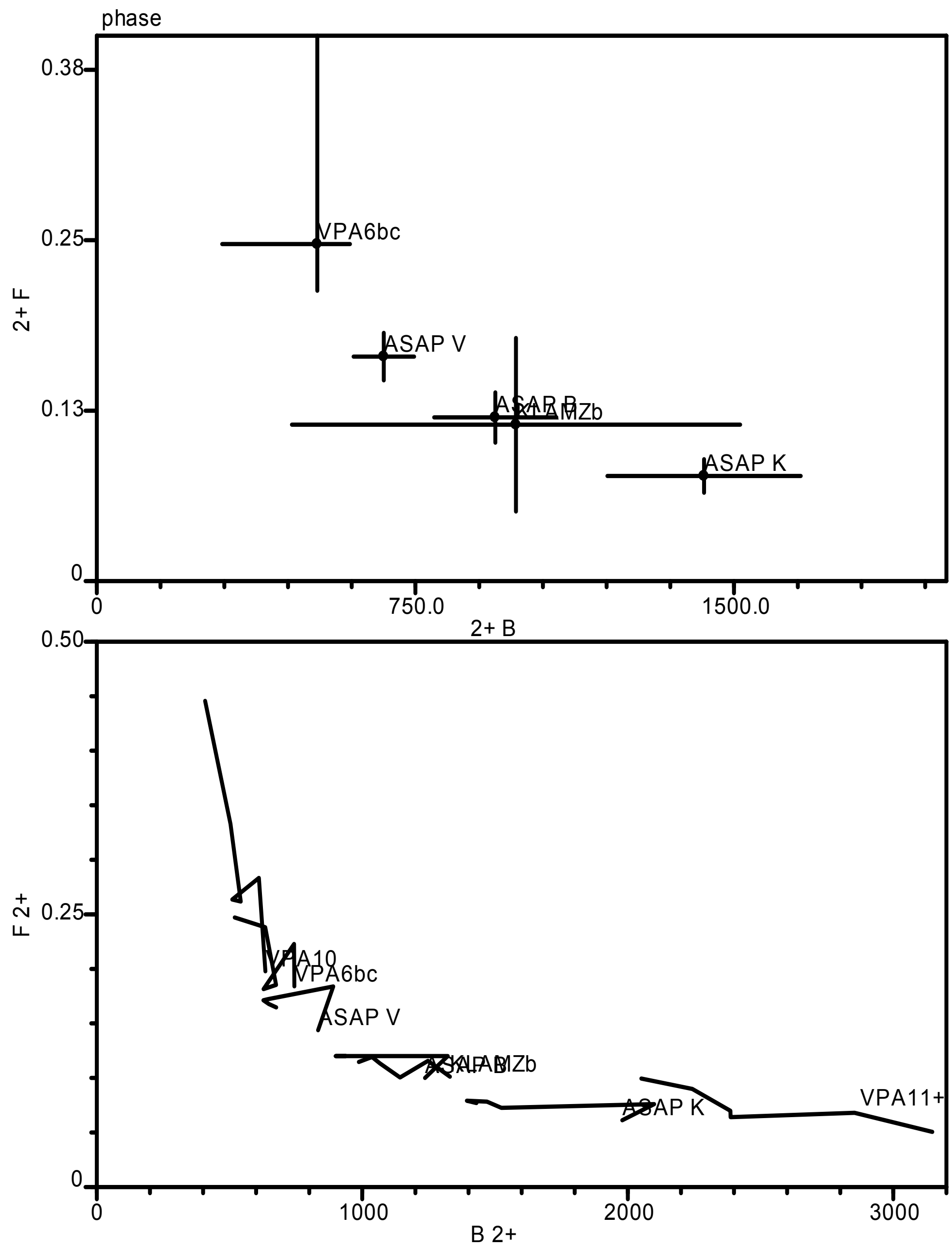

Figure 4. Summary plots of model runs. The upper panel is the terminal Biomass –Fishing Mortality from ASAP base (ASAP B), ASAP VPA-like (ASAP V), ASAP KLAMZ-like (ASAP K), a bias corrected VPA (VPAbc) and KLAMZ base (KLAZb). The error bars are single standard deviations from the Hessian matrix for all except VPAbc which is bootstrapped. The lower panel is the trajectory for the last 6 years for Age 2+ for VPA models which estimated numbers at age out to 11 in two different manners.

BENCHMARK FORMULATION

Rapporteur: K. Piner

Model Selection

Model selection was dependent upon the relative merits of the catch at age data relative to the indices. Because both data sources were considered informative but uncertain, an approach that uses both with appropriate weights (representing uncertainty) was deemed necessary. It was noted that whatever modeling approach is adopted, it will be in place for a long period of time and should therefore be robust to potential changes in the data.

The ASAP VPA – like and VPA models used aging information to determine year-class strengths whereas the ASAP KLAMZ – like and KLAMZ models required imposition of some structure e.g. stock / recruitment relationship. While the aging information for ages 6 and older was considered uncertain, that for the younger age groups was felt to be useful. There was general agreement that this information should be included in the assessment. This narrowed the choice amongst the models to the ASAP Base, ASAP VPA – like and VPA models. Of these, the ASAP Base model showed the least retrospective pattern and was considered to be compromise amongst all the formulations. It was thus chosen as the benchmark formulation. However, in drafting harvest advice, it will be important to emphasize that model selection was difficult and that there are considerable uncertainties in the determination of stock status. The recommendation was made to examine potential bias in this model using procedures such as MCMC.

The inclusion of the acoustic survey information was discussed, particularly its potential contribution to the retrospective pattern. Some work with VPA indicated that there was only a small improvement in the retrospective pattern by removing the acoustic survey time series. Although there was discussion as to whether the series should be removed from the benchmark formulation, the consensus was that there is no rationale to exclude it (contrary to the situation with the DFO survey) and thus it should be kept as one likelihood component with further examination undertaken as warranted.

It was noted that the weights at age used in the ASAP model are the commercial weights at age and not for the start of the year. The consensus was to use weights at age for the start of the year to calculate biomass.

There was consensus that the weighting of data sources used in the model is appropriate.

It was recommended that future assessments provide a measure of the statistical fit to the various data sources in addition to the residual plots.

Model Formulation

An ASAP is to be used using the annual catch at age, $C_{a,t}$ for ages a = 2 to 5, 6+ and time t = 1967 to the terminal year, where t represents the beginning of the time interval during which the catch was taken.

The ASAP is calibrated to the following survey indices:

NMFS spring, ages a = 2 to 5, 6+, time t = 1968 to 1984
NMFS spring, ages a = 2 to 5, 6+, time t = 1985 to terminal time
NMFS fall, ages a = 2 to 5, 6+, time t = 1967.5 to 1984.5
NMFS fall, ages a = 2 to 5, 6+, time t = 1985.5 to terminal time
NMFS winter, ages a = 2 to 5, 6+, time t = 1992 to terminal time
NMFS Acoustic, a=2 to 5, 6+, time t= 1999 to terminal time.

The lognormal errors in the indices are assumed independent and identically distributed. The relationship between the indices and population abundance is assumed proportional. The tuning indices' observations are all equally weighted.

The stock/recruitment relationship is turned off by estimating a mean level of recruitment and deviations about the mean level.

The numbers at age in the first year, 1967, are freely estimated.

Fishery selectivity is set to one for ages 2+.

Natural mortality is assumed to be 0.2 for all ages and years.

Maturity at age for all years is set to 0.21, 0.86, 0.93, 0.98, and 1.00 for ages 2 through 6+, respectively. Spawning stock biomass is calculated as the sum of mid-year numbers at age times maturity at age times weight at age. Total biomass is calculated outside the model using weights at age for the start of the year times numbers at age at the start of the year.

Harvest Reference Points

The Committee indicated that the reference points for consideration should be total biomass, fishing mortality, Bmsy, MSY and Fmsy and F proxies. Because ASAP has a flat S/R curve, MSY calculation based upon the S/R curve are not reliable and proxies may be needed. Given the similarity of the Fox and Schaefer surplus production estimates, which were performed external to the model, only the Fox estimates need to be presented. Furthermore, it was recommended that the calculation of population reference points use the population weights at age and fishery reference points use the fishery weights at age.

Using these weights at age, the Fox surplus production model reported earlier was updated. It estimated Fmsy = 0.31, MSY = 194,000 mt and Bmsy = 629,000 mt. Yield per recruit reference points (proxies for Fmsy) were estimated as F0.1=0.21 and F40%=0.20.

Projection Procedure

Only short term (two year) projections should be done using the starting population number at age provided by the benchmark formulation. The default population and fishery weights at age would be the arithmetic average of the last three years of the data set. Partial recruitment would be considered flat-topped for Age 2+. Natural mortality would be set at 0.2.

The Committee discussed if any adjustments should be made in the projections to account for the retrospective pattern. The consensus was not to explicitly account for the retrospective bias in the calculations and inform managers of the effects that the retrospective pattern might have on the projections.

The Committee discussed the utility of the risk analyses and concluded that the retrospective pattern was too problematic to undertake and include risk analyses in the Transboundary Status Reports (TSR).

Other Considerations

The Committee discussed three diagnostics that would indicate the need for a formal re-evaluation of the current assessment:

1. Poor standard model diagnostics
2. Sudden change in survey estimates
3. Industry not able to catch quota

These were not considered exclusive but would be good indicators to monitor.

The format of the TSR was also discussed. A proposal was made to include a new Sources of Uncertainty section in the TSR template in which all the sources of uncertainty would be consolidated. However, there is uncertainty associated with each of the elements of the benchmark and consequently it was felt more appropriate to discuss this in the sections in which it is relevant.

BIOMASS DISTRIBUTION

Overholtz, W. 2006. Relative Proportions of the Inshore and Offshore Components of the Gulf of Maine – Georges Bank Herring Complex. TRAC Working Paper 2006/01.

Rapporteur: K. Bolles

Presentation Highlights and Discussion (W. Overholtz)

There was a brief discussion on how the proportion information would be applied. The proportions would be applied to the total biomass produced by the assessment to provide the biomass by component. Then, the biomass by management area (Figure 1) would be calculated based upon a table within the current management plan showing the component composition by area.

It was asked if the three approaches (commercial acoustics, morphometrics, and survey ratios) all considered the proportions by component consistently by either numbers or biomass. It was replied that except of the morphometric analysis, where the number of samples is the metric, the others considered proportion in terms of biomass.

There then was discussion on the catchability assumptions of the acoustic surveys in the assessment models. Taking a ratio of the acoustic estimate to the total stock biomass implies use of a catchability of 1.0 or at least close to this. It was reported that the catchability in the KLAMZ model for the acoustic survey was first estimated at 0.91 and might on further analysis drop to 0.85. It was suggested that the decision on the utility of the approach await the final assessment results. [Note: this decision was revisited when the assessment formulation was confirmed and it was considered to proceed only with the commercial acoustic, morphometric and bottom trawl survey estimates of the proportions]

There followed an extended discussion on the morphometric study as well as questions on the limited tagging results.

It was asked if all three estimates of the proportions were equally valid and if so what is the most appropriate means to average them. It was agreed that they are equally valid and that the overall average be based on the average of each estimate. This gives equal weight to each method used to determine the proportion.

CONCLUDING REMARKS

The Co-Chairs thanked the Committee for its efforts to create a new benchmark formulation for the assessment of Gulf of Maine and Georges Bank herring. While it was recognized that there was considerable uncertainty in the agreed-to formulation that could only be resolved through further research, it was chosen as a compromise that took into account these sources of uncertainty. Notwithstanding this, a procedure was recommended that will guide future TRACs on herring until the results of the new research become available.

REFERENCES

Cadigan, N.G., and P.J. Farrell. 2005. Local Influence Diagnostics for the Retrospective Problem in Sequential Population Analysis . ICES Journal of Marine Science, 62: 256-265.

Gavaris, S. 2003. Transboundary Resource Assessment Committee (TRAC): Report of Meeting Held 10–14 February 2003. CSAS Proceedings Series 2003/014.

Legault, C.M., and V.R. Restrepo. 1999. A Flexible Forward Age-Structured Assessment Program. Int. Comm. Cons. Atl. Tunas, Coll. Vol. Sci. Pap. 49(2): 246-253.

Mohn, R. 1999. The Retrospective Problem in Sequential Population Analysis: An Investigation Using Cod Fishery and Simulated Data. ICES Journal of Marine Science, 56: 473-488.

Overholtz, W.J., L.D. Jacobson, G.D. Melvin, M. Cieri, M. Power, D. Libby, and K. Clark. 2004. Stock Assessment of the Gulf of Maine - Georges Bank Atlantic Herring Complex, 2003. Northeast Fisheries Science Center Reference Document 04-06. 290 pp.

APPENDICES

Appendix 1 : List of Participants

Participant	Affiliation/Address	Telephone	FAX	E-Mail
Karen Bolles	NOAA/NEFSC, USA	(774) 268-9031		Karen.bolles@noaa.gov
Noel Cadigan	NWAFC/DFO, Canada	(709) 772-5028		cadigann@dfo-mpo.gc.ca
Matt Cieri	ME/DMR, USA	(207) 633-9520		Matthew.cieri@marine.gov
Bob Cochrane	FWFA, Canada	(506) 755-6644		fundyweir@nbnet.nb.ca
Dave Ellenton	Cape Seafoods Inc. Gloucester, MA, USA	(978) 283-8522	(978) 283-3133	dave@capeseafoods.com
Stratis Gavaris	SABS/DFO, Canada	(506) 529-5912		gavariss@mar.dfo-mpo.gc.ca
Joe Grist	ASMFC/Science, USA	(202) 289-6400		jgrist@asmfc.org
Tony Hooper	Connors Brothers, Canada	(506) 456-1520		Tony.hooper@connors.ca
Larry Jacobson	NOAA/NEFSC, USA	(508) 495-2317		Larry.jacobson@noaa.gov
Chris Legault	NOAA/NEFSC, USA	(508) 495-2025		Chris.legault@noaa.gov
David Libby	ME/DMR, USA	(207) 633-9532		David.a.libby@maine.gov
Jay Lugar	Herring Science Council, Canada	(902) 492-2469	(902) 492-2293	logix@hfx.eastlink.ca
Ralph Mayo	NOAA/NEFSC, USA	(508) 495-2310		Ralph.mayo@noaa.gov
Claire MacDonald	FM/DFO, Canada	(902) 426-9854		macdonaldcm@mar.dfo-mpo.gc.ca
Clare McBane	NHF&G, Region 3 225 Maine Street Durham, NH 03824 USA	(603) 868-1095	(603) 868-3305	cmcbane@nhfgd.org
Gary Melvin	SABS/DFO, Canada	(506) 529-8854		melving@mar.dfo-mpo.gc.ca
Bill Michaels	NOAA/NEFSC, USA			William.michaels@noaa.gov
Robert Mohn	BIO/DFO, Canada	(902) 426-4592		mohnr@mar.dfo-mpo.gc.ca
Peter Moore	APA/New Bedford, USA	(207) 233-4353		petejmoore@aol.com
Robert O'Boyle (Co-Chair)	BIO/DFO, Canada	(902) 426-3526	(902) 426-5435	oboyle@mar.dfo-mpo.gc.ca
William Overholtz (Co-Chair)	NOAA/NMFS, USA	(508) 495-2256		William.overholtz@noaa.gov
Kevin Piner	SWFSC/La Jolla, USA	(858) 546-5613		Kevin.piner@noaa.gov
Michael Power	SABS/DFO, Canada	(506) 529-5881		powermj@mar.dfo-mpo.gc.ca
Rob Stephenson	SABS/DFO, Canada	(506) 529-5860		stephensonr@mar.dfo-mpo.gc.ca
Roger Stirling	SPANS, Canada	(902) 463-7790	(902) 469-8294	spans@ns.sympatico.ca
Mark Terceiro	NOAA/NEFSC, USA	(508) 495-2203		Mark.terceiro@noaa.gov
Mary-Beth Tooley	ECPA, USA	(207) 837-3537		ecpa@adelphia.net
Michele Traver	NOAA/NEFSC, USA	(508) 495-2195		Michele.traver@noaa.gov

Appendix 2. Meeting Terms of Reference

Background

The TRAC was established in 1998 to peer review assessments of transboundary resources in the Georges Bank area and thus to ensure that the management efforts of both Canada and USA, pursued either independently or cooperatively, are founded on a common understanding of resource status.

Prior to 2003, scientists from both countries had participated in each other's peer review of the Gulf of Maine herring assessments but there has not been a joint peer review meeting. During the 10 - 14 February 2003 meeting, the TRAC considered the assessment framework for the Gulf of Maine herring assessments. At that meeting, consensus was reached on how to deal with the stock complex and management units. It was deemed necessary to undertake an evaluation of the entire complex with subsequent consideration of the individual components. Evaluation of the relative proportions of the biomass between the inshore Gulf of Maine and Georges Bank should be considered to give guidance for the individual components. It was also established that surveys indices from the NMFS bottom trawl survey would be divided into and treated as two distinct time periods pre-1985 and 1985- present. Despite considerable investigation of model formulations for the assessment, a number of issues remained to be resolved. It was noted that to verify and compare models, new and revised data need to be considered. Thus a continuation of the hydroacoustic survey is likely to elucidate trends in biomass. Improved age determinations for older fish should give a better indication of the total mortality in the stock.

The purpose of this meeting is to review and incorporate any new information from survey indices and the fisheries, revisit the model formulation issues and recommend a suitable approach upon which to base management advice.

Objectives

- Review progress made on the recommendations from the 10 - 14 February 2003 TRAC meeting.
- Update results with the latest information from fisheries and research surveys.
- Review the assessment model formulation issues and recommend an approach for stock status determination.
- Apply the agreed assessment approach to update the status of the coastal stock complex of Atlantic herring through 2005 and characterize the uncertainty of estimates.
- Using available data from acoustic, trawl surveys, and recent tagging studies, evaluate the relative proportions of the biomass between the inshore Gulf of Maine and Georges Bank to give guidance on the usefulness and the degree to which these results can be used to provide advice to managers.
- Review the harvest strategy biological reference points to meet management requirements of both countries.
- Review approach for the provision of projections to meet the requirements of both countries

Products

- Meeting Proceedings, which will document the details of the review and summarize the consensus results.

Participation

- NEFSC and DFO Stock Assessment teams and other laboratory scientists
- Invited external (not from NEFSC or DFO Scotia Fundy) reviewers
- Representatives from US and Canadian management agencies
- US State and Canadian Provincial representatives
- US and Canadian fishing industry participants

Appendix 3. Meeting Agenda

1 May 2006 – Monday

Travel

2 May 2006 – Tuesday

09:00 - 09:30	Welcome and Introduction (Chair)
09:30 - 10:00	Overview of Issues and Recommendations from February 2003 TRAC Meeting
10:00 - 12:00	Review of Data Inputs with focus on the surveys
12:00 – 13:00	Lunch
13:00 – 16:00	Presentation & Discussion of Model Formulations
16:00 – 17:00	Plenary on meeting work plan

3 May 2006 - Wednesday

08:30 – 10:00	Presentation & Discussion of Biomass Distribution Calculation
10:00 – 12:00	Breakout
12:00 – 13:00	Lunch
13:00 - 15:00	Plenary on progress on analyses
15:00 - 17:00	Breakout

4 May 2006 - Thursday

08:30 – 10:00	Plenary & Consensus on Biomass Distribution Calculation & Model Formulations
10:00 – 12:00	Plenary Discussion on Reference Points & Projection Methods
12:00 – 13:00	Lunch
13:00 – 17:00	Breakout to undertake assessment

5 May 2006 - Friday

09:00 – 12:00	Review of Stock Status
12:00	Adjournment

Appendix 4. Observations on the TRAC Meeting and Process by N. Cadigan, R. Mohn, and K. Piner.

Introduction

Although reports such as these tend to focus on the negatives, overall the meeting was productive. It appears that the meeting was able to bridge the conflict among models that were developed in previous meetings. The TRAC was well chaired and the participants were engaged and constructive. The herring assessment is quite challenging compared to say traditional groundfish assessments, mainly because of the questions regarding the data.

Time limitation was the most pressing problem. If the week had been fully devoted to the definition of a benchmark assessment formulation, there still would not have been enough time. When the need to produce an assessment and an advisory document as well were added to the agenda, the time for full technical evaluation of relevant data, models and diagnostics became clearly impossible.

Procedural Considerations

There is a need to provide guidelines of expectations for the external reviewers. This report could have a precedent setting role. For example, should externals be required to act as rapporteurs? Are they expected to do any analysis? Are reviewers to participate in developing models or act strictly as independent peer review of the meeting proceedings?

In addition, because an advisory document was to be produced, some industry representatives participated in the meeting. It is difficult for these representatives to understand and follow the technical discussions involved in a benchmark assessment. Benchmark determination should be separated from the assessment to produce an advisory document.

Clearly, there was not enough time during the meeting to fully investigate either the data inputs or the model performance. Although the preferred model appears to be a reasonable representation of the stock dynamics and should be useful for management, there remain major uncertainties regarding the model performance that were not addressed during the meeting. Some of these are issues (convergence, bias, etc.) need to be addressed before we can feel fully comfortable with the assessment and subsequent results.

The time limitations were exacerbated by inadequate pre-meeting information distributed to the reviewers. A more clearly developed and timely process for disseminating information (possibly including data) to all participants would be beneficial to the process.

Many problems related to the herring assessment cannot be resolved during a benchmark meeting. There is a need for intercessional research to address problems identified at previous benchmark meetings e.g. aging, retrospective patterns, trends in survey catchability over time etc. Many of the sources of uncertainty need to be quantified (catch-at-age, stock composition - an important management issue). Is the design of the survey and analysis of catches suitable for pelagics? Perhaps intercessional generic workshops on inputs, say landings and sampling, would be a benefit across many stocks and represent an efficient approach to stock assessment.

Technical Considerations

As well as model selection, the selection of data sources did not receive sufficient attention. For example, during the write-up on Friday, there was still uncertainty about which sources of information were included in the models, and why they were included.

Adequate model diagnostics were not provided; nor was there time to review the diagnostics that were provided. The implication, or better, the cause of the large retrospective pattern in most model formulation requires more attention. Specifically, how to do projections when such a pattern is evident was and remains an issue.

More use of sensitivity analysis to data sources and models is required. For example the sensitivity of the potential impact of the apparent aging discrepancies would have been useful.

Is M at 0.2 for all ages and time reasonable? An M contour produced during the session suggested that it could be higher. Also, results from a multispecies analysis suggested that it has been time variable. Meta-analysis and some simpler methods (Hoenig's, Pauli's etc) may prove useful.

Decomposition of the likelihood (or any fit statistic) into the contribution of individual data sources would aid the understanding of model performance.

More information about commercial data would be useful. For example, trends in catches, effort and catch rates, spatial changes in the fishery, gear changes, etc. Commercial catches represent a lot of potential 'sampling'. Perhaps an industry survey could be contemplated. Also, information of this sort would be useful to link the early fishing effort and mortality to the recent period.

The assessment would benefit from a more complete explanation of uncertainty and how it is characterized and incorporated into the assessment and projections.

N. Cadegan
R. Mohn
K. Piner

Appendix 5. Landings, t, by Country and Area of Gulf of Maine – Georges Bank Herring Stock Complex.

YEAR	FOR GB[1]	U.S. GB[2]	CAN GB	FOR GOM	US GOM[3]	SNE[4]	MAT[5]	TOT US	NB[6]	TOTAL FOR	TOTAL
1960					60,237	261	152	60,650	34,304	34,304	94,954
1961	67,550	105			25,548	197	101	25,951	8,054	75,604	101,555
1962	152,141	101			69,980	131	98	70,310	20,698	172,839	243,149
1963	97,646	322			67,736	195	78	68,331	29,366	127,012	195,343
1964	130,949	489			27,226	200	148	28,063	29,432	160,381	188,444
1965	41,691	1,191			34,104	303	208	35,806	33,460	75,151	110,957
1966	138,396	4,308			29,167	3,185	176	36,836	35,805	174,201	211,037
1967	217,532	1,211		5,226	30,191	247	524	32,173	30,032	252,790	284,963
1968	372,840	758		21,497	40,928	245	122	42,053	33,145	427,482	469,535
1969	307,080	3,678		25,084	28,336	2,104	193	34,311	26,539	358,703	393,014
1970	245,283	2,011		13,716	28,070	1,037	189	31,307	15,840	274,839	306,146
1971	263,525	3,822		19,498	32,631	1,318	1,151	38,922	12,660	295,683	334,605
1972	171,408	2,782		24,220	37,444	2,310	409	42,945	32,699	228,327	271,272
1973	197,708	4,627		10,725	21,767	4,249	233	30,876	19,935	228,368	259,244
1974	146,155	3,370		7,865	29,491	2,918	200	35,979	20,602	174,622	210,601
1975	141,513	4,583		5,249	31,938	4,119	117	40,757	30,819	177,581	218,338
1976	42,758	744		921	49,887	191	57	50,879	29,206	72,885	123,764
1977	1,776	381		382	50,348	301	33	51,063	23,487	25,645	76,708
1978		2,059			48,734	1,730	46	52,569	38,842	38,842	91,411
1979		1,270			63,492	1,341	31	66,134	37,828	37,828	103,962
1980		1,700			82,244	1,200	21	85,165	13,525	13,525	98,690
1981		672			64,324	749	16	65,761	19,080	19,080	84,841
1982		1,378			32,157	1,394	20	34,949	25,963	25,963	60,912
1983		58			24,824	72	21	24,975	11,383	11,383	36,358
1984		53			33,958	79	10	34,100	8,698	8,698	42,798
1985		316			27,157	196	13	27,682	27,863	27,863	55,545
1986		586			27,942	632	20	29,180	27,883	27,883	57,063
1987		11			39,970	376	87	40,444	27,320	27,320	67,764
1988					39,568	1,307	365	41,240	33,421	33,421	74,661
1989					52,774	269	39	53,082	44,112	44,112	97,194
1990			91		54,192	761	48	55,001	38,778	38,869	93,870
1991			64		50,984	3,947	402	55,333	24,576	24,640	79,973
1992					55,948	716	4,564	61,228	31,968	31,968	93,196
1993					53,929	1,829	1,347	57,105	31,572	31,572	88,677
1994		474	266		51,413	1,935	502	54,324	22,241	22,507	76,831
1995		64			69,989	14,630	856	85,539	18,248	18,248	103,787
1996		1,758	2,491		78,885	26,876	1,079	108,598	15,913	18,404	127,002
1997		6,262	79		71,395	20,914	527	99,098	20,552	20,631	119,729
1998		31,067			52,683	20,084	1,903	105,737	20,092	20,092	125,829
1999		6,243			76,861	21,528	1,028	105,659	18,592	18,592	124,251
2000		16,171	275		64,839	27,275	568	108,853	16,830	17,105	125,958
2001	1,241	34,510	3,317		55,815	17,691	420	108,436	20,210	24,768	133,204
2002		14,325	1,605		67,523	10,557	753	93,157	11,807	13,412	106,569
2003		20,263			64,514	15,447	491	100,715	9,003	9,003	109,718
2004		8,952			73,800	11,658	30	94,439	20,620	20,620	115,059
2005*		13,397			65,273	14,310	278	93,259	12,639	12,639	105,898