Fisheries and Oceans Canada / Pêches et Océans Canada

NOAA FISHERIES
NATIONAL MARINE FISHERIES SERVICE

CERT

Comité d'évaluation des ressources transfrontalières

Comptes rendus 2005/01

TRAC

Transboundary Resource Assessment Committee

Proceedings 2005/01

Proceedings of the Transboundary Resources Assessment Committee (TRAC) Benchmark Review of Stock Assessment Models for the Georges Bank Yellowtail Flounder Stock

Woods Hole, Massachusetts
25 – 26 January 2005
and
26 – 29 April 2005

S. Gavaris[1], R. O'Boyle[2], and W. Overholtz[3]
TRAC Co-Chairs

[1]Fisheries and Oceans Canada
St. Andrews Biological Station
St. Andrews, New Brunswick
Canada

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

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

June 2005

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 during 25-26 January 2005 to discuss data issues, and also met during 26-29 April 2005 to complete a benchmark review of stock assessment models for the Georges Bank yellowtail flounder stock. These meetings were held to address the retrospective pattern that has been apparent in assessment of this stock during the past several years.

RÉSUMÉ

Le Comité d'évaluation des ressources transfrontalières (CERT) s'est réuni à Woods Hole les 25 et 26 janvier 2005 afin de discuter de questions de données. Il s'est réuni à nouveau du 26 au 29 avril 2005 pour procéder à un examen des points de référence des modèles d'évaluation de stock applicables à la limande à queue jaune du banc Georges. Il s'agissait dans les deux cas d'élucider la tendance rétrospective qui est apparue dans l'évaluation de ce stock au cours des quelques dernières années.

INTRODUCTION

At each meeting, the Co-Chairs welcomed meeting participants (Appendix 1), reviewed the agenda (Appendix 2), briefly mentioned the set of working papers to be presented, and reviewed the procedures for preparation of the meeting report.

At each meeting, the terms of reference (TORs) for the meeting (Appendix 3) were reviewed. The focus of the January meeting was to review the sources of data to be used in the benchmark assessment and produce estimates of other important sources of information (discards, etc.) that were not included in previous assessments. The April meeting evaluated the results of several working papers (Appendix 4), which investigated trends from survey indices, alternative model formulations, causes of retrospective patterns in this stock, and solutions for the retrospective problem. Various problems with the assessment were thoroughly discussed during the meeting. Research recommendations for further work on the Georges Bank yellowtail flounder stock were also outlined (Appendix 5).

DATA COMPILATION MEETING (25 – 26 January 2005)

Management Unit

The management unit currently recognized by Canada and the USA for the transboundary Georges Bank yellowtail flounder stock includes the entire Bank east of the Great South Channel to the Northeast Peak, which encompasses Canadian fisheries statistical areas 5Zj, 5Zm, 5Zn and 5Zh and U.S. statistical reporting areas 522, 525, 551, 552, 561 and 562. Based on new information from the comprehensive review by Cadrin (2003) and recent results from cooperative science/industry tagging programs conducted by Canada and the USA, there does not appear to be any justification for redefining the geographic boundaries of the stock management unit. This conclusion was accepted without discussion.

Landings

The USA fishery for yellowtail flounder on Georges Bank started in the mid-1930s, while the Canadian fishery began in 1993. Other countries harvested yellowtail flounder on Georges Bank during the 1960s and 1970s. No catches by countries other than the USA and Canada have occurred since 1975.

Canada

The majority of Canadian landings are made by otter trawl and the fishery generally occurs in the third quarter. Flatfish landed as unspecified flounder have been significant in some previous years and these have been apportioned to yellowtail flounder using the species composition of the identified flatfish.

Small amounts of landings attributed to longline gear are considered to be erroneous. The gear may have been incorrectly coded, but it is more likely that the species was coded incorrectly. Accordingly, these landings were not included in the assessment. Their exclusion should not have any noticeable effect because the amounts are very small.

The sex composition of the fishery landings has varied from year to year. The male to female ratio observed in DFO surveys has been somewhat higher than 50%, but has been variable and without trend. The variation of sex ratio in the fishery is likely due to the temporal and spatial patchiness of fishery operations and the underlying distribution of sexes, and is therefore not considered indicative of the sex ratio in the population.

USA

The principal fishing gear used in the USA fishery is the otter trawl, but sea scallop dredges and sink gillnets contribute some landings. Trawlers that land yellowtail flounder generally target multiple species on the Southwest Part of the Bank, on the Northern Edge, and along the western and southern boundaries of Closed Area II.

<u>Discards</u>

The principal sources of discards are considered to be the Canadian sea scallop fishery, the USA trawl fishery and the USA sea scallop fishery. Discards are not reported in Canadian fishery statistics but are reported in USA Vessel Trip Reports (VTR) for recent years. At-sea observations and/or VTR records can be used to determine discard rates and discards. Three common approaches for estimating discards are based on prorating observed or VTR discards based on the ratio of:

- discarded yellowtail to landed yellowtail
- discarded yellowtail to effort
- discarded yellowtail to landed scallops

While consistency in methodology is an important consideration, the specific circumstances and the nature of the available data determined the particular ratio applied in each situation, or if an alternative approach was used to obtain the most reliable results.

For an assessment to be meaningful, the inclusion of discards must be consistent over the entire time series. Discards cannot be included in some years and left out in others. Accordingly, attempts were made to derive discard estimates for all years. The reliability of discard estimates from all three principal sources suffered due to insufficient at-sea monitoring. It was therefore necessary to invoke unverifiable assumptions about patterns in discard rates when information and data were lacking. In late 2004, increases occurred in observer coverage of the sea scallop fleets in both Canada and the USA which in the future should allow direct estimation of discards without invoking unverifiable assumptions.

Canada

Yellowtail flounder discards in the Canadian sea scallop fishery have not previously been included in the assessment. Based on results from at-sea observer monitoring in 2001-2002, these discards are considered to be substantial relative to current landings. Any bycatch that is not landed (*i.e.,* not recorded in fishery statistics) was designated as "discards". The analysis was done separately for two periods. The period 1996-2003, when landing of yellowtail flounder was not permitted, was considered first because there was higher observer coverage. The period 1973-1995, when landing of yellowtail flounder was permitted, is based on very limited observer coverage. At-sea monitoring prior to 2001 was very sparse, compromising the reliability of discard estimates. The utility of these discard estimates for assessment analysis and their impact on the assessment results will need to be evaluated.

For the 1996-2003 period, when landing of yellowtail flounder was not permitted, effort in the sea scallop fishery was prorated by the observed discard rate of yellowtail to effort to obtain an estimate of yellowtail discards. Available data do not support any unit area trends in yellowtail discard rates. There was a tendency for higher discard rates of yellowtail flounder in April, May and June, and for lower discard rates in November and December. While the dispersion of observations was great, this was consistent with seasonal patterns of landings per unit of effort observed during 1986-1995. Exploratory analyses using discard rates pooled over quarters gave similar results indicating that the discard estimates were not sensitive to this assumption. Quarterly discard rates for periods lacking observed trips were derived by interpolation and application of a seasonal pattern. To estimate discards for year 1996 and later, quarterly discard rates were applied to the total quarterly effort of the scallop fleet.

During 1973-1995, the number of observed trips was very limited and hence the discard ratios subject to influence by anomalous outliers. Therefore, it was considered prudent to compare results from all three approaches. The discard estimates based on the discard to landed ratio did not appear consistent with the amount of effort. Discard estimates based on discard to effort ratios and discard to landed scallop ratios were fairly comparable. The landed scallop proration appeared somewhat susceptible to variation in scallop abundance, *e.g.* discards declined when scallop catch declined, while effort remained high and yellowtail abundance was thought to have remained relatively stable. Of the three approaches, the effort proration results appeared most consistent and were adopted. The seasonal factors applied in the 1996-2003 period were not used because these refinements were considered unwarranted given the other limitations of the available information for this period.

USA

Age specific discards for 1973-1993 were derived using (in order of preference and based on availability of data) sea sampling data, interviewed trips, or the difference in the catch-at-size between survey and landings subsequently smoothed using a retention model. The latter approach assumes that discarding occurred due to culling for minimum size. This assumption is considered appropriate for the groundfish trawl fishery. Comparison of size composition data of landed and discarded yellowtail flounder from the limited observed scallop fishing trips also supports this assumption for the scallop dredge fishery. This methodology and the results have been previously reviewed and were incorporated as reported. During the SAW 18 review, age 1 results were not reported because age 1 discards during 1988-1993 were not deemed to be reliably estimated. Subsequent assessments, however, have used these discard estimates and reported age 1 results. It is recommended that the analyses used to derive discards during 1973-1993 be documented.

Since 1994, discards have been estimated using fishery specific (trawl or dredge) half-year discard to landed ratios. Comparisons of discard rates derived using sea sampling data and Vessel Trip Reports (VTRs) suggested that VTR discard estimates were not significantly underestimated if only records that reported discards of any species were used. Trawl discard estimates were generated from VTR discard to kept ratios for the years 1994-1999 and from observer discard to kept ratios for the years 2000-2003 when there were sufficient observed trips. Dredge discards were estimated using three different estimation procedures: (a) for 1994-1998, dredge discard estimates were based on VTR discarded to landed ratios. The resulting discard estimates are small due to the low level of landings during this period; (b) for 1999 and 2000, special access programs to Closed Area II had high levels of observer coverage and generated high discard estimates; and (c) for 2001-2003, very few VTR or observer records were available for discard estimation. Additionally, no yellowtail were landed by the Georges Bank sea scallop dredge fishery in 2003 preventing the use of the usual discard to kept ratio. A regression of yellowtail discard estimates on sea scallop landings for the years 1994-2000 produced a tight relationship (r^2=0.98), which was used to estimate discards for the years 2001-2003 (during which scallop landings were within the range observed in 1994-2000). This regression contained an intercept term. While there are theoretical reasons not to use an intercept term, there were documented observer trips that caught scallops but not yellowtail - which could generate a negative intercept. Forcing the regression through the origin resulted in very minor changes to the estimates of dredge discards during 2001-2003 indicating that the estimates are not sensitive to the assumption made regarding an intercept in the regression. A second approach that corroborated these estimates used a single yellowtail discard to kept scallop ratio from VTR records for the years 2001-2003, which was applied to scallop landings in these years. This second approach produced discard estimates very similar to those from the regression approach.

Catch at Age

In past assessments, concern was expressed about the absence of older yellowtail flounder. Examination of consistency tests for age interpretation and of trends in average length at age over time did not reveal any patterns that might be indicative of drifting in age interpretation.

The traditional method for deriving catch at age by applying the respective age length key (ALK) to the length composition is the preferred approach. When the ALK sample size at each length is low, it may be necessary to “borrow” from other suitable ALKs to adequately determine catch at age. However, a low ALK sample size should not be a problem if most fish measured for the respective length composition were also aged.

An ALK (the proportion at age for any given length) is affected by year class strength of the cohorts in the population. Therefore, to apply the traditional method, the "borrowed" ALK data must come from the same year and from a season and area that display similar fish growth patterns.

If it is necessary to "borrow" ALK data from other years, an alternative method for deriving catch at age is required. One approach is to apply an appropriate growth template or length at age key (LAK) to the length composition using a numerical iterative technique (iALK). In essence, the iALK method starts with an initial guess of the catch at age, uses that to convert the LAK to an ALK, and then applies the ALK to the length composition to get an updated catch at age. When the comparison of the updated catch at age to the catch at age from the previous iteration meets a tolerance criterion, the final catch at age is accepted. The acceptability of iALK results hinges on obtaining an appropriate growth template.

Canada

Yellowtail landings in the Canadian fishery during 1993-2003 were processed by sex. While this protocol differs from that used for the USA fishery, it was considered appropriate if there were sufficient ages in sexed ALKs. There was very good sampling of the length composition in the Canadian fishery, but ALKs were not available. It was thus considered best practice to pool all available ALK information from USA port sampling, USA sea sampling and NMFS surveys to compile half-year ALKs that could be applied to the length composition data from the Canadian fishery. While the ALK sample sizes for these keys were substantially greater than those used before, the changes in catch at age were nominal. However, the ALK sample size for 1995 was only marginally acceptable. No Canadian fishery sampling data are available for years prior to 1993 when landings were low. Catch-at-age values for Canadian landings prior to 1993 were therefore derived (for both sexes combined) by multiplying the proportion of Canadian landings to USA landings by the USA fishery numbers-at-age.

Very limited data exist on the length composition of yellowtail discards in the scallop fishery. Length composition data from observed scallop trips in 2001, 2002 and 2004 was compared to survey length data to derive half year selectivity patterns. These selectivity patterns were applied to the annual survey length compositions to derive the length composition of the discards. Combined sex half-year ALKs generated from available data from USA port samples, USA sea samples and NMFS surveys were applied to the discard length compositions and then prorated to the total estimated discards to derive discards-at-age.

To avoid future "borrowing" of ALKs, NMFS has offered to age yellowtail samples from the Canadian fisheries.

USA

Sampling for length composition and ALKs was poor in the mid/late 1990s, but is otherwise considered adequate. Sea sampling data cannot be used to augment port sampling for length composition because landings are classified by size, while sea sampling catches are not sorted by market category. In any event, sea sampling was low when port sampling was poor. The 2001-2003 catch at age was redone with “borrowed” ALKs to boost sample size, but little impact was noted on the resulting catch-at-age.

Length frequencies available from the USA sea sampling program were considered sufficient to age the yellowtail discards in both the USA trawl and dredge fisheries in all years 1994-2003, except for the scallop dredge fishery in 2001. In this case, a selectivity ogive was developed using data from the NEFSC sea scallop survey and the 1999 special access program. This selectivity ogive was applied to the 2001 scallop survey to estimate the length frequency distribution of the yellowtail discards in the 2001 dredge fishery. ALKs were used in all cases to transform the length frequencies by half year to age frequencies.

Indices of Abundance

As with catch at age, an important consideration for indices of abundance at age is the transformation of length composition to age composition. A further consideration in analyzing survey data is alternative estimators. The delta estimator was examined using the NMFS surveys.

DFO Survey

Age determinations from DFO surveys were not available. NMFS spring survey ALKs, augmented with first half USA port sampling and sea sampling ALKs, were used to derive DFO survey abundance-at–age indices. This was considered the best practice with the available data. To avoid “borrowing”, NMFS has offered to age yellowtail samples collected on future DFO surveys.

NMFS Surveys

While the sample sizes of ALKs in some years are low, NMFS survey age compositions are considered reliable because the ALKs are dedicated keys and almost all fish measured for length composition were aged. A possible exception might be the reliability of older ages in early years of the survey.

The delta estimator was examined and the resulting time trends were almost identical to the traditional stratified means. Estimates of precision were also similar. Accordingly, the stratified mean estimates were retained.

Weight at Age

Fishery

Weight-at-age values for Canadian and USA landings and discards were derived using the applicable ALKs, LFs, and length-weight relationships. These values were combined to derive an overall fishery weight-at-age, weighting by the respective catch-at-age. A noted limitation is that the 1973-1993 weight-at-age matrix for USA discards is the same as for the landings.

Population

Fishery weight-at-age values were interpolated to beginning-of-year values using the method of Rivard (1980), and then used to calculate beginning of year biomass. To avoid fishery selectivity effects, average weights at age derived from NMFS spring surveys were used for ages 1 (mean = 0.012 kg, *n*=10) and 2 (mean=0.227 kg, *n*=13), where weights have been reported from NMFS spring surveys from 1992 to present.

ASSESSMENT METHODS MEETING (26 – 29 APRIL 2005)

Overview of Data and Issues for Benchmark

Stone, H. and C. Legault. A Comparison of Georges Bank Yellowtail Flounder VPA Results Using the 2004 Assessment Input Data and the Revised Input Data from the Benchmark Assessment Review. TRAC Working Paper 2005/01

Presentation Highlights

As part of the Georges Bank yellowtail flounder Benchmark Assessment Review, several changes were made to the indices and catch at age. The revised DFO indices showed some differences from age-specific indices used in past assessments, but were considered to provide better representation based on the age at length. The catch at age now includes discards of yellowtail from the Canadian offshore sea scallop fishery for 1973-2003 plus revisions to the USA discard estimates from offshore scallop and bottom trawl fisheries. It was considered best practice to pool all available ALK information from USA port sampling, USA sea sampling and NMFS surveys to compile half-year ALKs that could be applied to the length composition from the Canadian fishery and discards from the offshore scallop fishery. While the ALK sample sizes for these keys are substantially greater than what was used before, the changes in catch-at-age were nominal. The addition of the Canadian discards resulted in a slight increase in total catch over the time period used for the analytical assessment, 1973-2003. The associated revisions to mean weight at age were minor, generally occurred at random, and are not considered to have much influence on assessment calculations.

The VPA (ADAPT) results from the 2004 assessment as described by Legault and Stone (2004), were compared with those generated using the revised input data to provide an evaluation of the impact of the data changes and a starting point for exploring alternative assessment approaches. The VPA results using revised input data are not much different from the VPA results using 2004 input data. The observed differences (i.e. slightly higher recruitment at age 1, higher 3+ biomass and lower 4+ F in recent years) are mainly due to the addition of Canadian discard data in the catch at age. With the revised input data, there continues to be a strong retrospective pattern on age 4-5 F, spawning stock biomass and age 1 recruitment, with trends identical to those observed during the 2004 assessment. Terminal year population biomass and recruitment estimates are lower while estimates of F are higher when updated with additional year's data.

Discussion

It was concluded that the revisions to the data were warranted. The revised data were accepted and the assessment using the revised data deemed the basis for comparisons with alternative model formulations.

Sinclair, A. Estimation of Abundance and Mortality from Survey Indices. Addendum to TRAC Working Paper 2005/01

Results from the NMFS spring and fall and DFO bottom trawl surveys on Georges Bank were analysed for trends in abundance and mortality of yellowtail flounder. The consistency of year-class estimates both within and among surveys was examined using a correlation analysis among all possible survey-age combinations. The results indicated high consistency in year-class abundance estimates within and between surveys. In subsequent analyses, this was substantiated by close agreement among surveys in estimates of relative year-class strength and mortality rates.

Combined survey indices of relative abundance of recruits (ages 1-3 in the NMFS fall survey and ages 2-4 in the NMFS spring and DFO surveys) and adult yellowtail flounder (ages 4-6 in the NMFS fall survey and ages 5-7 in the NMFS spring and DFO surveys) were estimated using a multiplicative analysis (Sinclair et al. 1998). There was an initial decline in relative year-class strength for both recruit and adult ages from the early 1960s to a minimum in the early 1980s. There was a large increase in abundance of adult yellowtail flounder for year classes between 1992 and 1995. This increase occurred despite below-average year-class abundance at the recruiting ages. An index of total mortality between these age groups was also calculated. This indicated very low total mortality for these year classes.

Estimates of total mortality of adult yellowtail (ages 3-7) were obtained from a modified catch-curve analysis (Sinclair, 2001). This indicated a decline in total mortality in the 1990s followed by an increase in recent years. Total mortality is estimated to be above 1.0 in the most recent period.

<u>Population Model Investigations</u>

NOTE: Several working papers investigated alternative population models. Some discussion of these papers follows the presentation highlights. However, there was considerable overlap in the issues addressed and the general discussions are summarized in a section at the end of all presentations.

Cadrin, S. Biomass Dynamics Modeling of Georges Bank Yellowtail Flounder: Past, Present and Future. TRAC Working Paper 2005/02

Presentation Highlights

Biomass estimates from ASPIC are similar to those from age-based analysis until 1995, but are much greater than age-based estimates in recent years. Previous applications of ASPIC to Georges Bank yellowtail flounder were reviewed to identify problematic aspects. Previous analyses showed a relatively consistent estimation of F_{MSY}, annual fishing mortality since the late 1960s, and annual stock size since the 1970s. Conversely, estimates of B_{MSY} and biomass and fishing mortality in the 1960s were not as stable. Sequential revisions of software, data and model settings were made to improve model performance. More stable and precise parameter estimates were obtained by extending the catch series back to 1935 and assuming that 1935 biomass was at the carrying capacity. Although survey indices of relative stock size begin in 1963, 1935-1962 stock biomass values were estimated using the virgin stock assumption (1935 biomass=K), the observed catches during 1935-1962, and the estimated productivity from the 1963-2003 stock dynamics. A problematic pattern of model residuals persisted in which the model estimated greater biomass than indicated by all surveys in the last year of the analysis, 2003. The problem of large negative residuals in 2003 was not resolved with these modeling or data revisions, indicating that the model was overestimating current biomass (even in a relative sense, e.g., B/B_{MSY}). Model residuals were also large from survey observations during the 1960s and 1970s, because the model implicitly considers the rapid changes in survey indices to result from observation error. Unfortunately, the ASPIC model results and the recent survey indices offer divergent perspectives on current stock status. Given the past productivity of the stock when catches were relatively low, the model expects the increase in biomass during the 1990s to continue.

Performance of ASPIC was also evaluated using simulated data. Results from ASPIC analyses of simulated data with similar age-based attributes to the historical stock development of Georges Bank yellowtail indicate that relative trends in stock size and F from ASPIC were accurate, but absolute values were systematically biased. The average bias of fishing mortality estimates was -31%, and the average bias of biomass estimates was +24%. Simulation results confirm general attributes of ASPIC for other stock assessment applications, in that relative stock size and F appear to be reliable, but absolute estimates are not. However, ASPIC performed reasonably well in recovering underlying production parameters for the simulated data.

Results from ASPIC agree with those from the VPA (ADAPT) for 1973 to 1995, but ASPIC generates much greater biomass estimates from 1995 onwards. Therefore, recent information on age structure and survey indices does not support the increase in stock size indicated by ASPIC. The disagreement with recent survey indices, divergence from age-based estimates, and simulation results suggest that ASPIC should not be used for TAC projections. Despite the limitations of ASPIC for providing absolute projections, it may be useful for monitoring stock trends and providing management advice.

Discussion

The group was concerned that the estimate of current biomass, substantially greater than Bmsy, was overly optimistic. Comparison with historical data indicates that current geographic distribution of the stock is more restricted than in former periods (*e.g.,* much fewer yellowtail on the Southwest Part of the Bank). However, the current age structure of the catch is similar to that in the 1960s (Brown and Hennemuth 1971). The different pattern of F over time is not consistent with the survey-based estimates, suggesting that these are catch-based, not survey-based patterns. The assumption of logistic growth disagrees with productivity implied by a Beverton-Holt stock-recruit curve. However, ASPIC v5 may not be as efficient for estimating Fox model parameters (Prager 2004).

Legault, C. Exploration of the Georges Bank Yellowtail Flounder Data Using ASAP. TRAC Working Paper 2005/03

Presentation Highlights

ASAP is the acronym for Age Structured Assessment Program, a flexible forward-solving model. The original application of ASAP to Georges Bank yellowtail flounder in 2003 produced results similar to the ASPIC run, even though it used only the data available for the VPA (years 1973-2002 ages 1-6+). It was noted that the predicted catch for the plus group in the recent years was much higher than that observed (11-14% vs 1-3%).

A wider range of selectivity patterns were explored using ASAP and produced a wide range of possible current status that nearly spanned the large difference produced by the VPA and ASPIC analyses. It was also demonstrated that the VPA could produce a similarly wide range of possible solutions depending upon the assumptions made regarding the selectivity of older ages. The reason for the focus on selectivity was due to a few lines of evidence pointing towards Closed Area II (CA II) as a haven for older yellowtail flounder, thereby producing a dome selectivity pattern for the fishery.

This evidence supporting the hypothesis of CA II causing a dome in the recent selectivity pattern was used to justify selection of an ASAP run that forced a flat-topped selection pattern for years 1973-1993 and allowed estimation of the selectivity pattern for years 1994-2003. The selectivity pattern for the second period was estimated to be a dome with full selectivity at age 4 and declining to approximately 40% selectivity at the 6+ group. Projections of this model assuming a return to flat-topped selectivity, due to the expected opening of CA II, produced a higher catch than that from the VPA, but still lower than the status quo catch.

Given this evolution of ASAP application to Georges Bank yellowtail flounder and the current expansion of catch and survey data to their full range (instead of using the 6+ group), an initial attempt was made to detect a dome in the data using cohort catch curves. Catch curves for total catch appear to have shifted the age of full selectivity from age 2 to 4 over the time period examined. The catch curves do not show a flattening of the slope at older ages in recent years as would be expected due to a dome selectivity pattern caused by CA II. However, total mortality estimates do exhibit a decrease in recent years, although they are still quite high (>0.7).

Catch curves from the NEFSC fall and spring surveys exhibit quite a bit of noise, due to the relatively low numbers of animals sampled, especially at older ages. The total mortality estimates are also highly variable, but do not show an indication of reduction in recent years. The catch curves for the DFO survey also show quite a bit of noise, but provide a more consistent estimate of total mortality than do the NEFSC surveys.

Comparison of the total mortality estimates from the four sources shows a high level of mortality throughout the time period (averaging around 1.0) with little indication of a reduction in recent years, with the possible exception of the total catch data. This finding is consistent with the VPA results, but inconsistent with the ASPIC analysis and the general agreement that strong management measures since 1994 by both US and Canada have substantially reduced fishing mortality rates. The simplest ASAP run predicted catches for older ages in recent years that were substantially higher than the observed catches. The initial population abundance was inconsistent with later years. Overall, the results were similar to ASPIC results with F low initially, rising to approximately 1.0 in 1975 and then declining to approximately 0.05 in recent years, in strong contrast to the total mortality estimates from the catch curve analyses.

Stone, H., and S. Gavaris. Exploratory VPA Results for Georges Bank Yellowtail Flounder using various Formulations of ADAPT Software. TRAC Working Paper 2005/03

Presentation Highlights

ADAPT was used to examine a number of exploratory model formulations. The VPA used the revised annual catch at age, $C_{a,t}$, for ages a = 1 to 12, and time t = 1973 to 2003, where t represents the beginning of the time interval during which the catch was taken. The VPA was calibrated to bottom trawl survey indices, $I_{s,a,t}$, for:

s = DFO spring, ages a = 1 to 6+, time t = 1987 to 2004 (revised)
s = NMFS spring (Yankee 36), ages a = 1 to 6+, time t = 1982 to 2004
s = NMFS fall, ages a = 1 to 6+, time t = 1973.5 to 2003.5

Initial exploration using ADAPT with each survey indicated somewhat different results but generally similar trends. The NMFS scallop survey (age 1 recruits) and the early time series for the NMFS spring survey with Yankee 41 bottom trawl were not included in these exploratory analyses since they do not have much influence on the results.

For the DFO and NMFS fall survey, the age 6+ index (ages 6-9 aggregated) was applied to ages 6-9 in the CAA, while for the NMFS spring survey, the 6+ index (age 6-8 aggregated) was applied to ages 6-8 in the CAA. Fewer ages were aggregated in the NMFS spring survey because there appeared to be fewer age 9 fish present over the time series.

In all formulations, population numbers were estimated for ages 2-8 at the beginning of the year after the terminal year catch at age. In one case, population numbers were estimated for ages 2-12 and also estimated on the oldest age group back to 1983. A summary of the eight exploratory formulations is given below:

Formulation	M	*q*	Survey Indices
1	0.2	constant	1-5, 6+
2	0.4	constant	1-5, 6+
3	0.4 for 1973-1994 and estimated for 1995-2003	constant	1-5, 6+
4	0.2	*q* break (1973-1994, 1995-2003)	1-5, 6+
5	0.4	*q* break (1973-1994, 1995-2003)	
6	0.2	constant	1, 2, 3+
7	0.4	constant	1,2, 3+
8*	0.4	constant	1-5, 6+

*population numbers on oldest age group estimated back to 1983

These exploratory investigations could be summarized as encompassing three hypotheses of population dynamics:

– the high indices at ages 2 and 3 since 1995 represent high abundance, but the fish subsequently died of natural causes
    - increase in M

– the high indices at ages 2 and 3 since 1995 were due to increased survey catchability
    - increase in q

– the high indices at ages 2 and 3 since 1995 represent high abundance and the fish didn’t die, but fishery and survey catchability decrease at older ages
    - domed PR and domed q at age patterns

Assessment diagnostics were not very informative for evaluating these competing hypotheses. Judgment, based on expert knowledge of the fishery, is required to identify options supported by the weight of evidence.

The formulations explored to this point assume that the error in the catch at age is negligible. A formulation of ADAPT was developed that allows for random error in the catch at age. Models that allow for error in catch at age, sometimes referred to as ‘statistical catch at age’ models, assume that the error is random. Systematic error situations (*e.g.,* uncertainty in accuracy or completeness of catch reporting) are not necessarily handled better by statistical catch at age models.

An important consequence of allowing for error in catch at age is that certain structural assumptions - in particular assumptions that F on the oldest age can be calculated as a function of F at younger ages - cannot readily be incorporated. These types of F calculations are typically done deterministically based on the assumption that catch is known.

To conduct an informative comparison of the impact of allowing for error in the catch at age, it is therefore necessary to compare results to a model that does not allow for error but estimates the abundance of all or most year classes. The results from model formulation 8 are an appropriate comparison. The results from the catch error formulation were similar to those from model formulation 8 suggesting that allowing for random error in the catch at age is not an influential feature, but estimating the abundance at older ages (equivalent to estimating partial recruitment, *e.g.,* domed for older ages) has considerable impact.

Discussion

The large changes in survey catchability at the younger ages might imply a curvilinear relationship between population abundance and survey index. Another possibility is an increase in natural mortality, M, but what would be the biological mechanism to explain this? It was noted that these changes would be less dramatic if M were higher, say 0.4. However, previous estimates of M based upon tagging and total mortality versus effort relationships indicate an M of 0.2.

Discussion continued on possible changes in the survey catchability. Did the change in the NEFSC survey trawl doors affect yellowtail catchability? This is likely not the cause as the effect is seen in the DFO survey as well. Could survey catchability have changed over time? There may be a relationship of *q* with abundance that would be consistent with this. This led to discussion on the distributional characteristics of yellowtail, which tend to be patchy by sex. Whether or not sex related processes have influenced survey catchability changes over time could be examined through analysis of sex ratio changes by age and year. Changes in growth over years were also proposed as a mechanism for systematic changes in survey catchability.

Jacobson, L. and A. Seaver. Retrospective Patterns in Yellowtail Flounder. TRAC Working Paper 2005/5

Presentation Highlights

ADAPT VPA results for Georges Bank yellowtail flounder (Legault and Stone 2004) suffer from retrospective bias in abundance, fishing mortality and biomass estimates. The problems may be due to effects of area closures in 1994, errors in catch data, age-reading errors, natural mortality assumptions, plus-group calculations in ADAPT VPA, or other factors. These factors may act alone or in concert and are likely to be difficult to distinguish. For example, fishery data collection procedures and the condition of the stock changed in 1994 when some prime fishing grounds were closed to fishing.

We investigated patterns in input data and modeling conventions that contribute to retrospective bias in ADAPT VPA stock assessment estimates. Our analytical methods are based on Cadigan and Farrell's (2005) "local influence" approach. Our approaches differ from that of Cadigan and Farrell (2005) in being more general and avoiding assumptions about linearity. Results indicate that effects of closed areas on the distribution of older fish or errors in recent catch data may be responsible for retrospective bias patterns.

Adjustments to Catch-at-Age

Retrospective bias could be substantially reduced by decreasing catch at most ages during 1984-1994 and by increasing catch at most ages during 1994-2001. Adjustments to catches at ages 5-6 had the most substantial effects on retrospective bias. Combining ages, the largest reduction in annual catch was -11% in 1988 and 1991 and the largest increase in annual catch was 17% in 1997.

Using adjusted catch-at-age, the model usually under predicted survey abundances at ages 2+ (negative residuals) after 1994. The most pronounced differences in estimates from the original and adjusted catch-at-age data were in fishing mortality estimates during the 1980's. In addition to reducing retrospective bias, adjustments to catch-at-age reduced overall variability in estimates for recent years.

Adjustments to M at Age

Retrospective bias could be substantially reduced by decreasing M at most ages during 1983-1994 and by increasing *M* at most ages afterwards. There was no marked tendency for average adjusted M to increase or decrease with age. Survey residual patterns improved somewhat but the model usually under predicted survey abundances at ages 2+ (negative residuals) after 1994. Adjusted natural mortality rates were implausibly high for most ages and on average for years after 1995.

Adjustments to Survey Data

Preliminary runs using all survey data and all survey time except NEFSC spring survey data for 1973-1981 resulted in adjusted data that gave erratic abundance and mortality estimates. Further work with survey data was abandoned.

Refinements – Year Effects

Additional work focused primarily on catch-at-age data because catch-at-age data gave the best results overall. Although goodness of fit to survey data decreased, adjusted data were generally plausible and individual adjustments were seldom implausibly large.

An additional run assumed year effects for catch at age data. With year effects, the total annual catch varies but the proportions of the total catch in each age class do not. Year effects consistently increased total catch after 1994 (by a maximum of about 2.7 times in 1998) and consistently decreased total catch in earlier years (to very low levels during 1998). Year effects adjustments decreased retrospective bias more effectively than adjusting individual catch-at-age observations. Residual patterns improved when catch at age were adjusted for year effects.

Our results show that retrospective bias can be minimized simultaneously in abundance, F and biomass. Cadigan and Farrell's (2004) results and our results show that minimizing retrospective bias in one type of estimate may not effectively reduce retrospective bias in the others.

Best Results

Based on preliminary results, we hypothesize that catch-at-age errors were not the primary cause of retrospective bias. The most plausible adjustments to input data from this analysis were year effects for catch at age data. Goodness of fit to survey data substantially decreased but residual patterns were better. Retrospective bias was substantially reduced. However, catches were adjusted to implausibly high or low levels in some years.

Age misclassification is probably not the primary cause of retrospective bias because results indicate that catches of the oldest ages was overestimated prior to 1994 and underestimated afterwards. Difficulty in ageing old fish would probably cause undercounting of older fish in commercial fishery age composition data for all years.

Errors in specifying natural mortality rates seem unlikely as an explanation for retrospective bias because adjusted estimates were implausibly high for some ages in some years. Moreover, results suggest that M was too high prior to 1994 and too low afterwards and there is no other evidence for a change in M at that time.

Closed Area Effects

We hypothesize that retrospective patterns are due primarily to closed areas and, in particular, Closed Area II which was instituted late in 1994. The closed areas changed the degree of overlap between the fishery, the bulk of the stock, and large/old fish in particular. Beginning in 1994, yellowtail abundance increased and relatively large/old fish grew common in prime habitat areas within Closed Area II. In contrast to closed areas, fish abundance is low in areas open to fishing and large fish are scarce. Surveys are carried out inside of closed areas but fishing is largely prohibited. Thus, trends in abundance measured by the surveys, particularly of large/old fish, are not reflected in commercial catch. Our algorithm increased catches or natural mortality in recent years to reduce retrospective bias so that fishery catch-at-age and survey trends were more consistent. In effect, the algorithm was trying to account for fish that should have been in the catch, based on the survey data and assumptions regarding fishery selectivity in recent years.

ADAPT VPA models that estimate asymptotic fishery selectivity in recent years are implausible. Fishery selectivity estimates for the original model and models fit to adjusted data sets indicate that commercial fishery selectivity was asymptotic in recent years. However, survey data consistently indicate that large/old fish are present primarily in closed areas.

ASPIC results support the hypothesis that closed areas cause bias in ADAPT VPA estimates. Retrospective bias was not a problem in ASPIC surplus production model results (Legault and Stone 2004). ASPIC is not an age-structured model and would likely be less affected by differences in age structure of the stock, as measured by surveys, and the fishery, as measured by fishery catch at age.

Brodziak, J. Some STATCAM Analyses of Georges Bank Yellowtail Flounder, 1935 – 2003. TRAC Working Paper 2005/06

Presentation Highlights

A total of 14 statistical catch-at-age models were evaluated covering three assessment time horizons. Models 1a and 1b were constructed for 1973-2003, the VPA assessment time period. Models 2a-2d were developed for 1963-2003, corresponding to the longest research survey time series. Models 3a-3h were developed for 1935-2003, the entire time period of catch observations. All models used the revised catch-at-age. Catch biomass estimates were available for 1935-2003. Total catch included the 1940s-1960s discard estimates from Lux (1969). All models used the revised survey time series from NEFSC and DFO research surveys. Catch and stock weights at age and maturity fraction at age were based on the most recent updates.

For each of the models reported, fishery selectivity was estimated using two or more periods depending on model configuration. The baseline fishery selectivity periods were prior to 1994 and 1994-2003. These periods approximate the period preceding and following the implementation of Closed Areas I and II on Georges Bank.

Table 4. Summary of [illegible] models for Georges Bank yellowtail flounder.

Model	RMSE Catch Biomass	RMSE Fall Survey Index	RMSE Spring Survey Index	RMSE DFO Survey Index	RMSE Fishery CPUE Index	Spawning Biomass (mt) in 2003	Fishing Mortality in 2003	Spawning Biomass Depletion Ratio 2003 to 1994	Mean Recruitment (age-1, 000s)	Comments
(1a) GBYT, plus-group at age-6. 1973-2003. Initial equilibrium with catch at 1968-1972 average. Flat-topped fishery selectivity prior to 1994. Freely estimated fishery selectivity during 1994-2003. Freely estimated survey selectivities	0.033	0.546	0.395	0.426		51021	0.20	10.9	34157	Similar to ADAPT VPA configuration, although all ages used for tuning. Reasonable fit to survey biomass indices. Survey selectivities and recent fishery selectivity are all strongly domed.
(1b) GBYT, plus-group at age-6. 1973-2003. Initial equilibrium with catch at 1968-1972 average. Flat-topped fishery selectivity prior to 1994. Freely estimated fishery selectivity during 1994-2003. Flat-topped survey selectivities	0.113	0.562	0.432	0.468		46004	0.17	17.2	27605	Similar to 1a but survey selectivities constrained to be flat-topped. Note exponential increase in SSB and poorer fit to catch biomass (RMSE). There is still a strong dome in recent fishery selectivity.
(2a) GBYT, plus-group at age-6. 1963-2003. Initial equilibrium with catch at 1958-1962 average. Freely estimated fishery and survey selectivities	0.115	0.561	0.417	0.513		44286	0.20	10.1	39541	Also similar to VPA and 1a but time series extended back to 1963. Reasonable fit to noisy survey indices but similar domed selectivity patterns. Note the optimistic retrospective pattern (Fig 4).
(2b) GBYT, plus-group at age-9. 1963-2003. Initial equilibrium with catch at 1958-1962 average. Freely estimated fishery and survey selectivities	0.013	0.664	0.449	0.535		75272	0.14	10.2	36573	Same as 2a but with age structure of 1 to 9-plus. Note the fit to catch biomass is substantially improved. Strong doming in recent fishery and survey selectivities. Exponential increase in SSB.
(2c) GBYT, plus-group at age-6. 1963-2003. Initial equilibrium with catch at 1958-1962 average. Flat-topped fishery selectivity prior to 1994. Freely estimated fishery selectivity during 1994-2003. Flat-topped survey selectivity.	0.149	0.682	0.671	0.933		6711	1.70	1.9	32928	Differs from 2a in that survey selectivities are constrained to be flat-topped. Note the overpredict-underpredict pattern in survey biomass index fits in the 1980s-2000s. Much lower SSB and much higher, seemingly non-credible F.
(2d) GBYT, plus-group at age-9. 1963-2003. Initial equilibrium with catch at 1958-1962 average. Flat-topped fishery selectivity prior to 1994. Freely estimated fishery selectivity during 1994-2003. Flat-topped survey selectivity.	0.168	1.588	0.979	1.045		527967	0.02	56.5	33401	Same as 2c but using 9 age groups. Recruitment increases exponentially in late 1990s. Consistent with scenario of a small heavily fished stock that had a substantial biomass increase in recent years but where scale of increase is poorly determined.

Model	RMSE Catch Biomass	RMSE Fall Survey Index	RMSE Spring Survey Index	RMSE DFO Survey Index	RMSE Fishery CPUE Index	Spawning Biomass (mt) in 2003	Fishing Mortality in 2003	Spawning Biomass Depletion Ratio 2003 to 1994	Mean Recruitment (age-1, 000s)	Comments
(3a) GBYT, plus-group at age-6. 1935-2003. Unfished initial equilbrium. Freely estimated fishery and survey selectivities.	0.011	0.554	0.500	0.634		29652	0.23	9.2	28622	Similar to 1a and 2a with time horizon going back to 1935. No dome in fishery selectivities. No dome in spring or DFO surveys. Longterm decline in SSB, with increasing SSB since 1995. Recruitment deviations estimated from 1958 onwards.
(3b) GBYT, plus-group at age-6. 1935-2003. Unfished initial equilbrium. Freely estimated fishery and survey selectivities. Includes CPUE index.	0.013	0.530	0.517	0.661	0.147	23133	0.29	7.2	24013	Same as 3a but includes CPUE index from Lux (1969), with similar fit and interpretation.
(3c) GBYT, plus-group at age-6. 1935-2003. Unfished initial equilbrium. Flat-topped fishery selectivity prior to 1994. Freely estimated fishery selectivity during 1994-2003. Flat-topped survey selectivity. Includes CPUE index.	0.013	0.526	0.533	0.682	0.150	20022	0.35	6.5	23564	Simliar to 2c. Constrain survey selectivity to be flat-topped. Similar over/under pattern in survey index residuals. Regardless, fit is similar to 3a and 3b.
(3d) GBYT, plus-group at age-9. 1935-2003. Unfished initial equilbrium. Freely estimated fishery and survey selectivities.	0.080	0.601	0.423	0.519		62630	0.17	9.5	31381	Same as 3a but using 9 age groups. Strongly domed selectivities. Greater exponential increase in SSB.
(3e) GBYT, plus-group at age-9. 1935-2003. Unfished initial equilbrium. Flat-topped fishery selectivity prior to 1994. Freely estimated fishery selectivity during 1994-2003. Flat-topped survey selectivity.	0.011	0.579	0.536	0.686		31910	0.25	8.1	27520	Same as 3d but constrain survey selectivities to be flat-topped. Results are similar to 3a,3b,3c.
(3f) GBYT, plus-group at age-9. 1935-2003. Unfished initial equilbrium. 5-year fishery selectivity periods. Flat-topped fishery selectivity prior to 1995. Freely estimated fishery selectivity during 1995-present. Flat-topped survey selectivity.	0.010	0.576	0.536	0.690		34225	0.22	9.1	27102	Similar to 3e. Expanded fishery selectivity periods to be sequential five-year blocks. Not much practical impact on results.
(3g) GBYT, plus-group at age-9. 1935-2003. Unfished initial equilbrium. Four fishery selectivity periods (1935, 1977, 1985, 1994). Freely estimated fishery selectivity during 1994-present. Flat-topped survey selectivity.	0.012	0.589	0.539	0.694		30782	0.26	7.9	27210	Change number of fishery selectivity periods to match major fishery management changes in 1977, 1985, 1994. Not much practical effect but note estimation uncertainty for recruitment during 1980s. Note retrospective pattern (Fig 15).
(3h) GBYT, plus-group at age-9. 1935-2003. Unfished initial equilbrium. Flat-topped fishery selectivity prior to 1994. Freely estimated fishery selectivity during 1994-2003. Flat-topped survey selectivity. NEFSC surveys omitted during mismarked trawl warp period.	0.011	0.562	0.553	0.755		22626	0.36	5.8	26814	Similar to 3e, except excludes NEFSC survey indices during mismarked trawl warp period 2000-2002. Excluding these data tends to reduce SSB and increase F.

1973-2003 Assessment Time Horizon

Two models were developed for the 1973-2003 time horizon. The age structure of the stock was assumed to be in equilibrium with an annual catch equal to the average total catch during 1968-1972 (18930 mt). The two models differed in their treatment of survey catchabilities. Model 1a had freely estimated survey catchabilities, freely estimated fishery selectivity during 1994-2003, and fishery selectivity constrained to be flat-topped prior to 1994. Model 1b had survey catchabilities constrained to be flat-topped, freely estimated fishery selectivity during 1994-2003, and fishery selectivity constrained to be flat-topped prior to 1994. Model 1a provided the better statistical fit to the catch biomass and survey biomass index series, as indicated by the lower RMSE values. Despite providing reasonable fits to the observed survey indices, the estimated survey catchabilities for Model 1a were strongly dome-shaped. In particular, the NEFSC fall survey was severely dome-shaped with relatively low probabilities of capture for older yellowtail. Further, there was a substantial optimistic retrospective pattern.

1963-2003 Assessment Time Horizon

Four models were developed for the 1963-2003 time horizon. The age structure of the stock was assumed to be in equilibrium with an annual catch of 7920 mt, the average during 1958-1962. The four models differed in the modeled age structure and the survey selectivity shape. Models 2a and 2c used six (6) age groups. Models 2b and 2d used nine (9) age groups. Models 2a and 2b produced seemingly credible results but had strongly domed catchabilities. Models 2c and 2d did not produce credible results.

1935-2003 Assessment Time Horizon

Eight models were developed for the 1935-2003 time horizon. These differed in the number of age groups modeled, the use of the Lux (1969) CPUE index, whether survey catchability was constrained to be flat-topped, the number of fishery catchability periods, and whether NEFSC trawl indices during the mismarked warp period were used. The stock was assumed to be in an unfished condition in 1935, with an equilibrium of zero catch. The models produced qualitatively similar results. The best supported model produced seemingly reasonable fits to the survey indices and catch biomasses but exhibited a strong retrospective pattern and had the same overpredict/underpredict for recent survey index values. The consistent difficulties in fitting survey index values in the 1980s and 1990s suggest that a change in catchability may have occurred during this period. Such a change might be related to the spatial distribution of the resource inside and outside of closed areas.

Discussion

The discussion on the STATCOM approach focused on the differences between it and the other models discussed at the meeting. STATCAM and ASAP are both forward projecting statistical catch-at-age frameworks. STATCAM allowed domed surveys catchability and recruitment deviations were from constant recruitment rather than deviations from a stock-recruit relationship.

Patterson, K. 2005. Yellowtail flounder in Georges Bank: Assessment issues. TRAC working paper 2005/7

Presentation Highlights

The following critical issues have been identified:

1. A strong retrospective pattern exists with the existing model structure.
2. There is a large residual pattern indicating a change after the early 1990s.
3. Although the estimated biomass trend increases, commercial CPUE decreases or at least does not increase.
4. There are many fewer older fish in the data than are predicted by the assessment models.
5. Total mortalities estimated from survey data indicate an increasing trend, but this trend is not seen in the assessments.
6. The latest survey data (in 2004) are not consistent with previous assessments.

Some model explorations were carried out using a separable model framework. The starting point was that data prior to the early 1990s are not comparable with later data, and should not be treated as part of the same time series. Attention was focused on the last years of the time series, being those most critical for future management decisions.

Within the period 1993-2004, poor model diagnostics persist: there is a trend in residuals, most marked for ages 2 and 3 of the NMFS surveys, and a substantial retrospective effect persists. It appeared problematic to achieve a model fit that replicated the large increase in survey abundances in 2001 followed by a rapid decrease: the modeled dynamics of the population did not change sufficiently rapidly to follow this feature.

Two further changes were introduced: firstly, a higher natural mortality rate was assumed (=0.4) based on analyses done at this meeting and because of the smaller age-range of the population compared to other species for which an assumption of M=0.4 is routinely made. This higher natural mortality corresponds to faster stock dynamics, so allowing the fitted populations better to track decreases in survey abundances.

Secondly, visual inspection of residual trends suggested a possible nonstationarity of catchability with respect to abundance. A power catchability model fitted to estimate such effects indicated a strong density-dependence in some cases, especially for age groups 2 and 3 in the NMFS survey, while in others the estimated exponent of the catchability function was close to 1 (*i.e.,* most other catchabilities were estimated to be linear). An appropriate power-model catchability function could not be estimated in fitting the scallop index. The model was then reformulated with a linear catchability function for only the scallop index, and tested by retrospective analysis. The outcome indicated a good stability of the estimates from 1999 onwards. Earlier periods could not be tested due to overparameterization problems.

To improve robustness to possible highly-variable data series, an inverse-variance reweighting procedure was implemented with the constraint of a maximum weight of 1.5. The effect of this was small, with most weights estimated in the range 0.9 to 1.5.

The model formulation introduced above appeared to perform well with respect to the problem issues previously identified:

1. The retrospective performance of the model appears good.
2. The post-1990s residual problem is addressed, by simply removing data during the earlier part of the time-series.
3. The trends of fitted populations in the last years now coincide with the survey-estimated mortality rates (increasing mortality) and with information from the fishery (declining stock sizes).
4. Abundances of older fish now appear adequately fitted by the model.
5. The model is also reasonably consistent with the 2004 survey observations. Although these observations still force a downwards revision to perceptions of stock size, this revision is of the order of only about 25%.

Introducing higher M, the power-model catchability at ages 2-3, the removal of the pre-1993 data and the inverse-variance reweighting appear to have resolved a number of problem areas. Although the underlying biological understanding of these effects can be modeled, they are not necessarily understood.

Not all changes have been tested separately. Changing to a power-model catchability alone (retaining the full time-series of data, M=0.2, uniform unit weighting) did not remove the retrospective pattern. This comparison suggests that addressing both the issue of the residual blocks in the early/mid 1990s, and the apparent nonlinearity of ages 2 and 3 subsequently is required to remove the retrospective inconsistencies.

The following issues, possibly important, have not yet been addressed:

- Parameter estimates may be sensitive to model formulations. Investigations have indicated that relatively small changes to model structures can have quite large effects on parameter estimates. Consequently, management procedures in place should be robust to such mis-specifications.

- An issue of the effect of the closed areas on selection pattern was identified whereby a closed area can result in an apparent dome-shaped selectivity pattern by the fishery, and therefore the introduction of a closed area would introduce a change in selection. This may introduce a structural bias when a separable model is used to represent the populations. An "ADAPT" formulation may be more robust to such effects by having weaker selection pattern constraints

Further Model Explorations

A discussion on the various topics raised in the working papers ensued. The focus of the discussion was on the retrospective pattern, the residual pattern which started in 1995, the lack of older fish in the survey, the observation that survey total mortality (Z) was not decreasing while fishing mortality (F) from the assessment was decreasing, and the change in survey trend over the last several years. Additional analyses to address these patterns were proposed and completed. Some involved investigation of data issues while others involved examination of processes within the population (*e.g.,* M), the survey (*e.g.*, catchability changes), and the fishery (*e.g.,* partial recruitment).

Data Issues

Catches were added to the VPA input from 1995 through 2003, ranging from -129 t (2003) to 1,621 t (2000), representing about 20-40% of the catch included in the base run. The retrospective pattern remained for biomass, while that for F was reduced, but the F was estimated to be much lower than in the base run. However, adding catch and lowering F does not seem reasonable. Also, the residual pattern after 1994 was not as pronounced in this run but was still problematic. A discussion ensued of the possibility of misallocation of catches from the Georges Bank to the Cape Cod stock, based on VTR proration. It was felt that this was not likely since catches from the Cape Cod stock are an order of magnitude lower than those from Georges Bank. Also the ‘missing catch’ is in the older ages, making it unlikely that a misallocation of catch across all ages could be responsible

Using ADAPT, a longer-term retrospective analysis (1973 – 2003) was completed which confirmed that the retrospective problem began in the mid 1990s. Before then, there were no consistent patterns.

Preliminary length data were examined from the 2004 Georges Bank Special Access Program (SAP). There are indications from US observer trips in 2004 of possible larger fish coming out of the Closed Area II SAP.

Population Processes

Natural mortality (M) changes were investigated, by age and year. VPA formulations for ages 1 to 12, with no temporal split in survey q, were attempted. The first formulation held M constant at 0.2, the next estimated one M for the entire time period (0.6), another split M before and after 1995 and estimated different M for younger and older ages, and finally M was estimated along groups of cohorts. The results showed recent low Fs with biomass increasing then declining in some cases in recent years. M estimates ranged from around 0.1 to almost 0.9 depending on the cohort. There was only a limited analysis of the consequent retrospective patterns. As well, for changes in M to be credible, some biological process needs to be hypothesized, which was not apparent.

Survey Processes

To investigate density dependent processes in survey catchability, a power catchability relationship was included in a VPA model. A curvilinear trend in q was observed for ages 2 and 3, but q was nearly linear for the older ages. The retrospective pattern in F and residual trends was much improved. The group concluded that by allowing for a power trend in q, one could reduce the retrospective pattern without having to assume either a change in M over time, or missing catch. It was noted that the power function could be a proxy for some other process. One explanation is density-dependent juvenile natural mortality, which could result from a predator effect on large year classes.

Diagnostic work was presented on the above formulation. Changing M and inverse variance weighting were not considered important, but dropping the pre-1993 survey data improved the retrospective pattern. Including the early data, even with a power function on the catchability, causes a severe retrospective pattern with a clear step in the mid-1990s. This step pattern may indicate that the stock size is responding to the decline in the survey indices.

A comparison was undertaken of the residuals from the base run with those using a power catchability relationship on ages 1-3. The resulting residual pattern since the mid 1990s was less pronounced but still existed. However, the main point of the analysis was to emphasize the need for objective criteria to select between the models, which had not yet been done.

ADAPT was run by splitting the survey q for entire time series between 1994 and 1995, which improved the residual patterns, and resulted in a much higher F in recent years. Inspection of weight-at-age values and survey q's from the split series suggested that survey q may have increased for ages 2-3 because of increased size at age. However, the split q also indicated much greater q at older ages post 1995, and the retrospective pattern still persisted in the base ADAPT when only post 1995 indices were used.

To examine whether or not the predicted survey indices fell within the observed variance, survey means and stratified variance were examined for trends and patterns with a comparison of the predicted aggregate index from ADAPT with the survey index and its variance. The predicted index was generally within the 10-90th percentile, except after the late 1990s when the predicted index was below the lower percentile. The group concluded that the perspective on the current status of the stock (biomass and F) depends on the most recent survey observations influencing the calibration of the VPA. The pattern is similar to the post-1994 pattern of high residuals.

The sensitivity of the modeling results to the younger age indices was examined. The indices were high at ages 2 and 3, and low at ages 4, 5, and 6-9. Reported catch and assumed M were not sufficient for the declines from age 3 to older ages. So the VPA was calibrated with only with ages 4, 5, 6 to 9, with no young age group indices included. Recruitment was estimated using the partial recruitment. The results still exhibited positive residuals flipping to negative after 1994. The results suggested that the biomass trend reflects recent survey declines, and F has been increasing since mid-1990s, reflecting the Z pattern from surveys. However, the residual pattern since the mid-1990s still showed many positive values. Our impression of a low F and increasing biomass is primarily coming from the high indices at ages 2 and 3.

Mean-variance relationships in surveys were analyzed. With the exception of a few points, there was no effect of the asynchronous trawl warps on the catches of yellowtail flounder in all strata (as well as the important strata 16 and 19). There was a substantial increase in the clumping factor in the spring survey between the pre-1994 and post-1994 periods, but not for the fall survey. A high clumping factor results from a high variance to mean relationship. Overall, the analysis indicated that yellowtail flounder are highly clumped, and that clumping in the spring survey increased after 1994.

A survey q split series model was also run using STATCAM, with the split occurring between 1994 and 1995. Survey catchabilities were not fixed at constant values. Pronounced residual patterning occurred in most surveys before and after 1995. Spawning biomass showed a very rapid increase with no recent declines, but the confidence intervals were large. A pronounced domed survey selectivity was present in most cases. There was some question as to whether q increased or decreased between the first and second period in the DFO survey. The retrospective pattern in this run was still evident but somewhat reduced.

Fishery Processes

Size-at-age by cohort from the surveys was examined. Sexual dimorphism in growth begins between ages 2 and 3, with female growth thereafter greater than males. This could affect the partial recruitment to the fishery after age 3, but was not explored further.

A sensitivity analysis assigning different values for survivors at age 13 was completed. Values ranged from 0.1 to 100,000 fish surviving to age 13. Results were not sensitive to values up to 1000 fish. Higher numbers gave higher recruitment and biomass estimates, and lower Fs. It was recommended that 1000 fish be used to start subsequent VPAs.

An analysis was completed to show how to a perceived dome in fishery partial recruitment could be generated from a closed area. A spreadsheet calculation of F was accomplished based on a combination of a small portion of the population remaining inside a closed area subject only to M and the larger portion of the population fished outside the closed area. Overall, the simulated population was dominated by the younger ages from the open area, but older ages dominated within the closed area. This generated a fishing mortality pattern with low F at older ages, producing a dome.

General Discussion

In all of the above analyses, an inconsistency was apparent between the data and the model dynamics beginning in the mid-1990s. Adding catch post-1994 reduced the retrospective pattern, but there is no rationale to do so based on any external criteria. Increasing M post-1994 contributed to a reduction in the retrospective pattern, but there is no information to support an increase in natural mortality post-1994. Changes in survey catchability also contributed to a reduction in the retrospective pattern. There is some evidence of density dependent *q* for ages 2 and 3. There may be some other spatial process related to the imposition of Closed Area II in 1994, but it is not obvious what the biological mechanism would be.

The available information was not sufficient to allow discrimination among the potential hypotheses of the causes of the retrospective and residual patterns. Most involve processes in the younger age groups, but research is required to resolve these issues. Until new information becomes available, the group discussed what could be done and thus returned to the model explorations discussed above.

The VPA variant, in which the younger age groups were omitted, was revisited. This formulation makes no assumptions on the processes occurring in the younger age groups, but simply excludes the calibration indices for the younger ages, thereby focusing the analysis on ages 4 and older. The model generates a biomass trajectory that follows the recent trend in the survey indices, and is appealing in that it removes some conflicting information (ages 2 and 3) rather than incorporating additional information (increased catches or an increase in M). The weakness of this approach is that it relies on constant fishery partial recruitment at ages 2 and 3 for recent years, and uses this PR pattern to estimate the incoming year classes. Two other formulations also appeared promising for addressing these young age related problems: (1) truncating the survey time series; and (2) using a power function for survey catchability at ages 2 and 3.

Additional suggestions included basing advice on: (a) the comparison of the three approaches; (b) a model which included many technical fixes; (c) a base model (with minor changes); and (d) a more data based approach. There was considerable discussion on exactly what the latter approach was. Use of data alone as the basis of assessment does not remove the need to make assumptions – these are just more implicit. The AIM (An Index Method) was tabled as an example of a data-based approach. This method is based on two primary metrics: the Relative F (landings / survey biomass index with a 2 year running average), and the Replacement Ratio (index in year t/moving average for year t-1 to t-5, *i.e.*, the current stock size relative to the parental stock size that produced it). A Replacement Ratio of 1.0 indicates that the population is replacing itself. The two processes are highly correlated. To account for this, a randomization test is performed to test whether the correlation is due to chance. If not, estimates of the Relative F at a Replacement Ratio of 1.0 are bootstrapped to derive a probability distribution. For Georges Bank yellowtail flounder, the results using this approach with the NEFSC spring surveys are similar to those from ASPIC (low F, high biomass). This approach is similar to some model-based approaches that de-emphasize some pieces of data to remove some of the tension (in this case, age structure). The problem with the post-1994 data in this model is reflected in the calculation of Relative F, as it is in the ASPIC model. Should the entire time series be used to calculate the Replacement Ratio if there has been a change in the way the data (catch or survey q) can be interpreted before and after 1994?

<u>Benchmark Formulations</u>

General Approach

The group discussed approaches for the June 2005 TRAC meeting, including criteria to guide formulation of the consensus assessment models. It was agreed that the 2004 base age-structured model should be adjusted to address the data issues identified at the present meeting and should incorporate the revised data. It was agreed that the benchmark assessment should strive to:

1. minimize the residual pattern
2. reduce the retrospective pattern
3. include the fullest range of data available
4. include no untenable assumptions
5. strive for parsimony in model structure and assumptions

The most urgent problems and potential solutions to guide the benchmark formulations for the June 2005 TRAC meeting included:

Problem 1: Large block of residuals identified in early-1990s indicating non-stationarity in relationship between surveys and populations reconstructed from catches.

Problem 2: Large retrospective pattern in recent years, F appears successively underestimated.

Problem 3: Since 1995, model estimates inconsistent with survey trends in total mortality and abundance

The following table indicates (√ in cell) what aspect of the problems that the various proposed changes to data and model assumption would address.

		Problem	
Change to Data or Model Assumption	**1**	**2**	**3**
Add assumed non-reported catch by age and year to reported post 1994 catch (up to 1600 tonnes in one year)	√	√	√
Set M at all ages=0.4	-	-	-
Adjust M by age by year	√	√	√
Truncate survey series prior to 1995	√	-	√
Fit separate survey catchabilities pre & post 1995	√	-	√
Fit power function catchabilities for age 2 & 3	-	√	√
Use survey indices for age 4+ only	-	-	√
Ignore age structure	√	√	-

Consensus Benchmark Formulations

No single formulation was considered fully adequate for determining stock status. The procedure recommended for determining stock status uses two VPA model formulations corresponding to a model with major and minor changes and trends derived from survey data.

Assessment results from analyses conducted in the most recent years have displayed retrospective patterns, residual patterns that are indicative of a discontinuity starting in

1995, and fishing mortality rates that are not consistent with the decline in abundance along cohorts that is evident in the survey data. Essentially, the catch at age data and assumed natural mortality cannot be reconciled with the high survey abundance at ages 2 and 3 and low survey abundance at ages 4 and older.

The empirical evidence suggests that significant modifications to population and fishery dynamics assumptions are required to reconcile the observations from the fishery and the survey. Models that adopt these modifications to assumptions imply major consequences on underlying processes and/or fishery monitoring procedures. The magnitudes of implied changes to the natural mortality rate, survey catchability relationships, and/or unreported catch are so great that it makes the acceptability of the models that incorporate these effects suspect.

In view of the reservations about the implications to underlying processes, adoption of a benchmark formulation that incorporates these modifications as the sole basis for management advice was not advocated. It was therefore recommended that management advice be formulated after considering the results from three (3) approaches:

1. projections from a model that adopts assumptions towards reconciling the fishery and survey observations;

2. projections from a base case model with due consideration for the magnitude of the retrospective; and

3. trends in relative abundance and relative mortality rates derived from survey and fishery data.

It was suggested that when the indications from these three approaches are not coherent, the conservation implications of taking action on the assumption that one is correct when another is closer to reality should be described.

The specifications of the formulations for each of the approaches are described below.

Approach 1: Model with Major Changes

A VPA using the revised annual catch at age (including US and Canadian discards), $C_{a,t}$, for ages a = 1 to 12, and time t = 1973 to the terminal year, where t represents the beginning of the time interval during which the catch was taken. The error in the catch at age was assumed to be negligible compared to the error in the survey indices. Natural mortality was assumed to be 0.2 for all ages and years.

The VPA is calibrated to bottom trawl survey indices, $I_{s,a,t}$, for:

s_1 = DFO spring, ages a = 2 to 5, 6-9, time t = 1987 to 1994
s_2 = DFO spring, ages a = 2 to 5, 6-9, time t = 1995 to terminal time
s_3 = NMFS spring (Yankee 41), ages a = 1 to 5, 6-9, time t = 1973 to 1981
s_4 = NMFS spring (Yankee 36), ages a = 1 to 5, 6-9, time t = 1982 to 1994

s_5 = NMFS spring (Yankee 36), ages a = 1 to 5, 6-9, time t = 1995 to terminal time
s_6 = NMFS fall, ages a = 1 to 5, 6-9, time t = 1973.5 to 1994.5
s_7 = NMFS fall, ages a = 1 to 5, 6-9, time t = 1995.5 to terminal time
s_8 = NMFS scallop, ages a = 1, time t = 1983.5 to 1994.5
s_9 = NMFS scallop, ages a = 1, time t = 1995.5 to terminal time

The aggregated ages 6-9 survey indices were compared to ages 6-9 population abundance. The error in the indices was assumed to be independent and identically distributed. The relationship between indices and population abundance for ages 4 and older was assumed to be proportional, while that for younger ages was assumed to be a power relationship. Population abundance at age 1 in the terminal year was assumed equal to the geometric mean over the most recent decade. Population abundance in the terminal year is estimated for all other ages where the results are deemed reliable and calculated based on average partial recruitment to the fishery for any remaining ages. The survivors at age 13 in all years are assumed to be few, e.g. 1,000 fish or less.

Approach 2: Model with Minor Change

A VPA using the revised annual catch at age (including US and Canadian discards), $C_{a,t}$, for ages a = 1 to 12, and time t = 1973 to the terminal year, where t represents the beginning of the time interval during which the catch was taken. The error in the catch at age was assumed to be negligible compared to the error in the survey indices. Natural mortality was assumed to be 0.2 for all ages and years.

The VPA is calibrated to bottom trawl survey indices, $I_{s,a,t}$, for:

s_1 = DFO spring, ages a = 4, 5, 6-9, time t = 1987 to terminal time
s_2 = NMFS spring (Yankee 41), ages a = 4, 5, 6-9, time t = 1973 to 1981
s_3 = NMFS spring (Yankee 36), ages a = 4, 5, 6-9, time t = 1982 to terminal time
s_4 = NMFS fall, ages a = 4, 5, 6-9, time t = 1973.5 to terminal time

The aggregated ages 6-9 survey indices will be compared to ages 6-9 population abundance. Errors in the indices are assumed to be independent and identically distributed. The relationship between indices and population abundance for all ages are assumed to be proportional. Population abundance at age 1 in the terminal year is assumed equal to the geometric mean over the most recent decade. Population abundance in the terminal year is estimated for ages 4 and older where the results are deemed reliable and calculated based on average partial recruitment to the fishery for any remaining ages. Abundance and age 2 and 3 in the terminal year will be based on the average partial recruitment to the fishery in previous years. The survivors at age 13 in all years are assumed to be few, e.g. 1,000 fish or less.

Approach 3: Survey Trends

Trends in survey indices and total mortality derived from surveys.

Reference Points

The group considered the implications of model changes on reference point calculations and agreed that the proposed model changes would not significantly alter reference point calculations and agreed that Fref=0.25 be retained.

To account for the retrospective pattern, it was suggested that managers could select an F below Fref in proportion to the long term retrospective bias that will ensure, with a degree of confidence, that Fref will not be exceeded. This is similar in concept to the application of precautionary approach reference points, an approach employed by the ICES Study Group on Management Strategies.

CONCLUDING REMARKS

The Co-Chairs thanked the committee for its efforts in attempting to create a new benchmark framework for Georges Bank yellowtail. While the group could not define a single model formulation that addressed all the issues pertaining to this assessment, this was not for lack of effort. An amazing array of in-depth analyses had been attempted on what proved to be a very resistant problem. More research is required to allow discrimination amongst the competing hypotheses, some of which are summarized in Appendix 5. Notwithstanding this, a procedure was recommended that will guide the June 2005 TRAC meeting and indeed future TRACs on yellowtail until the results of the new research are available.

REFERENCES

Brown, B.E., and R.C. Hennemuth. 1971. Assessment of the yellowtail flounder fishery in Subarea 5. ICNAF Res. DOC. 71/14. 57p.

Cadigan, N.G., and P.J. Farrell. 2005 . Local influence diagnostics for the retrospective problem in sequential population analysis. ICES Journal Of Marine Science; Vol. 62; Number 2; pp. 256-265; 2005.

Legault, C.M.; and H.H. Stone. 2004. Stock Assessment of Georges Bank (5Zhjmn) Yellowtail Flounder for 2004. TRAC Reference Document 2004/03. 102 p.

Lux, F.E. 1969. Landings per unit effort, age composition, and total mortality of yellowtail flounder,/ Limanda ferruginea/ (Storer), off New England. ICNAF Res. Bull. 6, 47-69.

Prager, M.H. 1994. A suite of extensions to a nonequilibrium surplus-production model. Fish. Bull. 92, 374-389.

Sinclair, A.F., Chouinard, G.A., and Currie, L. 1998. Assessment of the cod stock in the southern Gulf of St. Lawrence, January 1998. Canadian Stock Assessment Secretariat Research Document 98/08.

Sinclair, A.F. 2001. Natural mortality of cod (*Gadus morhua*) in the Southern Gulf of St. Lawrence. ICES J. mar. Sci. 58: 1-10.

APPENDICES

Appendix 1. List of Participants

25 – 26 January 2005

Participant	Affiliation/Address
W. Overholtz (co-chair)	NMFS, Woods Hole
C. Legault	NMFS, Woods Hole
S. Cadrin	NMFS, Woods Hole
L. Van Eeckhaute	DFO, SABS
S. Gavaris (co-chair)	DFO, SABS
H. Stone	DFO, SABS
J. Burnett	NMFS, Woods Hole

26 – 29 April 2005

Participant	Affiliation/Address
R. O'Boyle (co-chair)	DFO, Dartmouth, NS
W. Overholtz (co-chair)	NMFS, Woods Hole
C. Legault	NMFS, Woods Hole
S. Cadrin	NMFS, Woods Hole
G. Chouinard	DFO, GFC
L. Van Eeckhaute	DFO, SABS
S. Gavaris	DFO, SABS
S. Wigley	NMFS, Woods Hole
H. Stone	DFO, SABS
R. Mayo	NMFS, Woods Hole
J. Hansen	DFO, Dartmouth, NS
T. Nies	NEFMC
L. O'Brien	NMFS, Woods Hole
A. Sinclair	DFO, PBS
L. Col	NMFS, Woods Hole
M. Palmer	NMFS, Woods Hole
A. Richards	NMFS, Woods Hole
L. Jacobson	NMFS, Woods Hole
M. Traver	NMFS, Woods Hole
K. Sosebee	NMFS, Woods Hole
A. Seaver	NMFS, Woods Hole
Jon Brodziak	NMFS, Woods Hole

Appendix 2. Agenda

25 – 26 January 2005
No agenda required

26 – 29 April 2005
26 April – Tuesday
09:00 - 09:30 Welcome and Introduction (Chair)
09:30 - 11:00 Overview of Data & Issues for Benchmark
11:00 - 12:00 Working Papers on ADAPT and ASPIC with current data & results

12:00 – 13:00 Lunch

13:00 – 17:00 Working Papers on New Methods
13:00 - 14:30 Adapt and Aspic continued (if needed)
14:30 - 16:00 WP 1 on ASAP
16:00 - 17:30 WP 2 on ADAPT new approaches

27 April - Wednesday
09:00 - 10:30 WP 3 on New Method
10:30 - 12:00 WP 4 on New Method

12:00 – 13:00 Lunch

13:00 - 14:30 WP 5 on New Method
14:30 - 17:00 Work Time

28 April - Thursday
09:00 – 12:00 Review of Reformulations

12:00 – 13:00 Lunch

13:00 – 17:00 Discussion & Resolution of Framework

29 April - Friday
09:00 – 12:00 Meeting Report Preparation

12:00 – 13:00 Lunch

13:00 – 17:00 Meeting Report Review
17:00 Adjournment

Appendix 3: Terms of Reference

Background

Since 1998, the Transboundary Resources Assessment Committee (TRAC) has reviewed stock assessments and projections necessary to support management activities for shared resources across the USA Canada boundary in the Gulf of Maine-Georges Bank region. These assessments are necessary to advise decision makers on the status of these resources and likely consequences of policy choices. The TRAC employs a two-tiered review process in which each of the stocks periodically undergoes an intensive peer review of the assessment model and assumptions. This is termed a benchmark assessment review. The benchmark assessment framework is applied as required, generally on an annual schedule, to provide the peer reviewed assessment of the resource status to fisheries managers.

The scope of a benchmark assessment can be categorized under four broad themes, a) definition of management unit, b) best approach for estimation of current status, c) determination of harvest strategy reference points and d) establishing projection procedures for evaluation of management tactic options. In any benchmark, not all themes may be subjected to evaluation. The extent of evaluation of a theme is defined by the terms of reference of the benchmark assessment. However, the basis of accepted practice for all themes should be documented.

The need to review the yellowtail flounder benchmark assessment framework was raised during the June 2004 TRAC meeting. The review will focus on themes b and d, addressing theme c to the extent necessary as a result of model formulation changes and will simply document the basis for the established management unit, theme a. Two sets of meetings were deemed necessary to adequately review the framework, the first to consider the data input and summarization and the second to consider the most appropriate model formulation for synthesis of available information. These will be conducted during January – April 2005 in time to be applied for the June 2005 TRAC assessment review meeting.

Process

Data Compilation Meeting (25 - 26 Jan 2005, Woods Hole)

This meeting will review methods to process the data inputs for use in the framework synthesis. The Terms of Reference for the meeting are:

1. Review the documentation on rationale for the current management unit
2. Review fishery landings and estimates of discards from all fisheries
3. Determine most appropriate methods for calculating fishery landings at age
4. Determine most appropriate methods for calculating discards at age from all fisheries
5. Determine most appropriate methods for calculating weights at age for the fishery catch

6. Determine most appropriate methods for calculating survey indices of abundance at age
7. Determine most appropriate methods for calculating weights at age for the population
8. Explore fishery catch per unit effort indices for use as tuning indices
9. Examine tagging data for application in stock assessment
10. Set date for availability of finalized data to be used in benchmark assessment

Assessment Methods Meeting (25 – 29 April 2005, Woods Hole)

At this meeting, analytical procedures for synthesizing all available information to determine the current status of the resource will be reviewed. The Terms of Reference for this meeting are:

1. Explore full range of assessment methods for estimating current abundance and exploitation rate such as, but not limited to, catch curves, separable VPA, index based approaches, surplus production, delay-difference, calibrated VPA, errors in catch at age models
2. Determine assessment approach that will be used in next assessment
3. If required, update reference points for harvest strategy based on agreed assessment approach
4. Formulate projection procedures for harvest advice based on agreed assessment approach

Products

- Proceedings, which document the details of the assessment framework
- Reference Documents, which support the technical basis for the recommendations

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 4: List of Documents

ANON. 2005. Draft proceedings of the January 2005 meeting of the TRAC: Georges Bank yellowtail flounder benchmark meeting. 6p.

Brodziak, J. 2005. Some STATCAM analyses of Georges Bank yellowtail flounder, 1935-2003. TRAC working paper 2005/6. 41p.

Cadrin, S. 2005. Biomass dynamics modeling of Georges Bank yellowtail flounder: Past, present, and future. TRAC working paper 2005/2. 37 p.

Cadrin, S. 2005. ASPIC analyses of revised simulations for Georges Bank yellowtail flounder. TRAC working paper 2005/2 addendum. 7p.

Jacobson, L., and A. Seaver. 2005. Retrospective patterns in yellowtail flounder. TRAC working paper 2005/5. 20p.

Legault, C. 2005. Exploration of the Georges Bank yellowtail flounder data using ASAP. TRAC working paper 2005/3. 11p.

Patterson, K. 2005. Yellowtail flounder in Georges Bank: Assessment issues. TRAC working paper 2005/7. 4p.

Sinclair, A. 2005. Estimation of abundance and mortality from survey indices. TRAC working paper 2005/1 addendum. 4p.

Stone, H., and C. Legault. 2005. A comparison of Georges Bank yellowtail flounder VPA results using the 2004 assessment input data and revised input data from the benchmark assessment review. TRAC working paper 2005/1. 20p.

Stone, H., and S. Gavaris. 2005. Exploratory VPA results for Georges Bank yellowtail flounder using various formulations of ADAPT software. TRAC working paper 2005/4. 37p.

Stone, H., and S. Gavaris. 2005. Exploratory VPA results for Georges Bank yellowtail flounder using various formulations of ADAPT software. TRAC working paper 2005/4 addendum. 7p.

Appendix 5: Research Recommendations

Catch audit:

- Request report from those responsible for ensuring data of the highest quality.

M related research:

- Tagging studies using the Southern New England area in lieu of Georges Bank.
- Food habits from fish, not from seals or sea birds.

RV q and density dependent processes:

- Examine length-based catchability patterns.
- Examine proportion at age by sex, i.e. a change in the proportion of a given sex over time.

Spatial processes:

- Explore stock assessment models that include area effects, literature available for sedentary species, examine transfer rates, tagging data can yield migration rates out of closed areas.