Atlantic States Marine Fisheries Commission

1050 N. Highland Street • Suite 200A-N • Arlington, VA 22201
703.842.0740 • 703.842.0741 (fax) • www.asmfc.org

MEMORANDUM

July 22, 2012

To: **ASMFC Atlantic Herring Section**

From: **Mark Robson, Law Enforcement Committee Coordinator**

Subject: **Law Enforcement Committee comments on Addendum V**

Members of the Law Enforcement Committee (LEC) of the Atlantic States Marine Fisheries Commission reviewed the management options contained in Atlantic Herring Addendum V for consideration of any law enforcement concerns or issues. The review was based on 3 components of the addendum:

1) Refining the sampling protocol for spawning area closures.

2) Considering a mechanism to shift the boundary between areas based on sampling data.

3) Consolidating all regulations into one primary management document.

No concerns or issues were raised by LEC members regarding these management measures. Current management allows for changes in spawning area seasons depending on the availability of sampling data, and no additional problems are foreseen with a mechanism to change boundaries provided that timely notification of such changes is integral to the process. The LEC is supportive of efforts to consolidate and standardize regulations into one primary management document.

The LEC appreciates the opportunity to review this addendum and provide input.

The following pages contain abridged reports from the SAW 54 where both Atlantic herring and Yellowtail flounder underwent benchmark assessment reviews. The included reports contain all the information relevant to Atlantic herring, but information regarding Yellowtail flounder were removed. For full reports please go to the SAW 54 website: http://www.nefsc.noaa.gov/saw/reports.html

A summary of the assessment report and each of the peer reviewers reports follow. The full assessment report will not be available for about a month (beginning of September). Staff will email the full assessment report to the Section as soon as it is made available.

Northeast Fisheries Science Center Reference Document 12-14

54th Northeast Regional Stock Assessment Workshop (54th SAW)

Assessment Summary Report

by the Northeast Fisheries Science Center

July 2012

Northeast Fisheries Science Center Reference Document 12-14

54th Northeast Regional Stock Assessment Workshop (54th SAW)

Assessment Summary Report

by the Northeast Fisheries Science Center
NOAA National Marine Fisheries Service
Northeast Fisheries Science Center, 166 Water St., Woods Hole, MA 02543

US DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Northeast Fisheries Science Center
Woods Hole, Massachusetts

July 2012

Northeast Fisheries Science Center Reference Documents

This series is a secondary scientific series designed to assure the long-term documentation and to enable the timely transmission of research results by Center and/or non-Center researchers, where such results bear upon the research mission of the Center (see the outside back cover for the mission statement). These documents receive internal scientific review, and most receive copy editing. The National Marine Fisheries Service does not endorse any proprietary material, process, or product mentioned in these documents.

All documents issued in this series since April 2001, and several documents issued prior to that date, have been copublished in both paper and electronic versions. To access the electronic version of a document in this series, go to *http://www.nefsc.noaa.gov/nefsc/publications/*. The electronic version is available in PDF format to permit printing of a paper copy directly from the Internet. If you do not have Internet access, or if a desired document is one of the pre-April 2001 documents available only in the paper version, you can obtain a paper copy by contacting the senior Center author of the desired document. Refer to the title page of the document for the senior Center author's name and mailing address. If there is no Center author, or if there is corporate (*i.e.*, non-individualized) authorship, then contact the Center's Woods Hole Laboratory Library (166 Water St., Woods Hole, MA 02543-1026).

Editorial Treatment: To distribute this report quickly, it has not undergone the normal technical and copy editing by the Northeast Fisheries Science Center's (NEFSC's) Editorial Office as have most other issues in the NOAA Technical Memorandum NMFS-NE series. Other than the four covers and first two preliminary pages, all writing and editing have been performed by the authors listed within.

Information Quality Act Compliance: In accordance with section 515 of Public Law 106-554, the Northeast Regional Office completed both technical and policy reviews for this report. These predissemination reviews are on file at the Northeast Regional Office.

This document may be cited as:

Northeast Fisheries Science Center. 2012. 54th Northeast Regional Stock Assessment Workshop (54th SAW) Assessment Summary Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 12-14; 40 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://www.nefsc.noaa.gov/nefsc/publications/

Table of Contents

SAW-54 ASSESSMENT SUMMARY REPORT

Introduction

The 54th SAW Assessment Summary Report contains summary and detailed technical information on two stock assessments reviewed during June 5-9, 2012 at the Stock Assessment Workshop (SAW) by the 54th Stock Assessment Review Committee (SARC-54): Atlantic herring (*Clupea harengus*) and Southern New England Mid-Atlantic yellowtail flounder (*Pleuronectes ferrugineus*). The SARC-54 consisted of 3 external, independent reviewers appointed by the Center for Independent Experts [CIE], and an external SARC chairman from the NEFMC SSC. The SARC evaluated whether each Term of Reference (listed in the Appendix) was completed successfully based on whether the work provided a scientifically credible basis for developing fishery management advice. The reviewers' reports for SAW/SARC-54 are available at website: http://www.nefsc.noaa.gov/nefsc/saw/ under the heading "SARC 54 Panelist Reports".

An important aspect of any assessment is the determination of current stock status. The status of the stock relates to both the rate of removal of fish from the population – the exploitation rate – and the current stock size. The exploitation rate is the proportion of the stock alive at the beginning of the year that is caught during the year. When that proportion exceeds the amount specified in an overfishing definition, overfishing is occurring. Fishery removal rates are usually expressed in terms of the instantaneous fishing mortality rate, F, and the maximum removal rate is denoted as $F_{THRESHOLD.}$

Another important factor for classifying the status of a resource is the current stock level, for example, spawning stock biomass (SSB) or total stock biomass (TSB). Overfishing definitions, therefore, characteristically include specification of a minimum biomass threshold as well as a maximum fishing threshold. If the biomass of a stock falls below the biomass threshold ($B_{THRESHOLD}$) the stock is in an overfished condition. The Sustainable Fisheries Act mandates that a stock rebuilding plan be developed should this situation arise.

As there are two dimensions to stock status – the rate of removal and the biomass level – it is possible that a stock not currently subject to overfishing in terms of exploitation rates is in an overfished condition, that is, has a biomass level less than the threshold level. This may be due to heavy exploitation in the past, or a result of other factors such as unfavorable environmental conditions. In this case, future recruitment to the stock is very important and the probability of improvement may increase greatly by increasing the stock size. Conversely, fishing down a stock that is at a high biomass level should generally increase the long-term sustainable yield. Stocks under federal jurisdiction are managed on the basis of maximum sustainable yield (MSY). The biomass that produces this yield is called B_{MSY} and the fishing mortality rate that produces MSY is called F_{MSY}.

Given this, federally managed stocks under review are classified with respect to current overfishing definitions. A stock is overfished if its current biomass is below $B_{THRESHOLD}$ and overfishing is occurring if current F is greater than $F_{THRESHOLD.}$ The table below depicts status criteria.

		BIOMASS		
		B <$B_{THRESHOLD}$	$B_{THRESHOLD}$ < B < B_{MSY}	B > B_{MSY}
EXPLOITATION RATE	F>$F_{THRESHOLD}$	Overfished, overfishing is occurring; reduce F, adopt and follow rebuilding plan	Not overfished, overfishing is occurring; reduce F, rebuild stock	F = F_{TARGET} <= F_{MSY}
	F<$F_{THRESHOLD}$	Overfished, overfishing is not occurring; adopt and follow rebuilding plan	Not overfished, overfishing is not occurring; rebuild stock	F = F_{TARGET} <= F_{MSY}

Fisheries management may take into account scientific and management uncertainty, and overfishing guidelines often include a control rule in the overfishing definition. Generically, the control rules suggest actions at various levels of stock biomass and incorporate an assessment of risk, in that F targets are set so as to avoid exceeding F thresholds.

Outcome of Stock Assessment Review Meeting

Text in this section is based on SARC-54 Review Panel reports (available at http://www.nefsc.noaa.gov/nefsc/saw/ under the heading "SARC-54 Panelist Reports"). For **Atlantic herring**, the Panel accepted the new ASAP assessment model. A feature of this new model is the 50% increase in natural mortality rate (M) during 1996-2011. This new M estimate is consistent with data on consumption of herring by predators and it largely resolves the retrospective pattern which has been a prominent feature of previous assessment models. The biological reference points were derived assuming that the 50% increase in M due to herring consumption will continue over the next 3 – 5 years. This assumption about the future is a source of uncertainty. The new biomass reference points (B_{TARGET} and MSY) are much lower than those from the previous assessment. A source of uncertainty in the stock projections is the size of the 2009 age-1 recruitment, which has been estimated to be almost twice as large as the next largest recruitment (1994). The 2009 age-1 fish contribute to the recent increase in stock biomass, and are a significant component of projected yield to the fishery in the future. It will be important to monitor the size of this year-class. Overall, the Panel concluded that the Atlantic herring stock is not overfished and that overfishing is not occurring.

For **Southern New England Mid-Atlantic yellowtail flounder** the Panel accepted a new stock assessment model (ASAP). There was a significant revision of most of the assessment's data sets. The new model assumed a higher natural mortality rate (M). There has been a marked decline in recruitment since 1990. Two stock–recruitment scenarios were developed which account for this decline, and the two scenarios lead to very different conclusions about biomass stock status. A "recent recruitment" scenario assumes that incoming year-classes since 1990 have been weak, perhaps due to a reduction in stock productivity, and not related to SSB. Alternatively, a "two-stanza" scenario assumes that recruitment over the entire time series is a function of spawning stock biomass (SSB) and that below about 4300 mt SSB average recruitment is very low. While neither scenario could be ruled out, the Panel concluded that the evidence was 60:40 in favor of the "recent recruitment" scenario (i.e., productivity change). Overall, the fishing mortality (F_{MSY}) reference point is relatively certain, and overfishing is likely not occurring. However, the reference points associated with biomass ($B_{MSY,}$ MSY) are uncertain due to the productivity change issue and require further exploration. There is considerable uncertainty as to whether or not the stock is overfished. Under the "recent recruitment" scenario the stock would not be considered overfished and it would be considered rebuilt to a new, much lower biomass target. In contrast, under the "two-stanza" scenario the stock would still be considered overfished.

Glossary

ADAPT. A commonly used form of computer program used to optimally fit a Virtual Population Assessment (VPA) to abundance data.

ASAP. The Age Structured Assessment Program is an age-structured model that uses forward computations assuming separability of fishing mortality into year and age components to estimate population sizes given observed catches, catch-at-age, and indices of abundance. Discards can be treated explicitly. The separability assumption is relaxed by allowing for fleet-specific computations and by allowing the selectivity at age to change smoothly over time or in blocks of years. The software can also allow the catchability associated with each abundance index to vary smoothly with time. The problem's dimensions (number of ages, years, fleets and abundance indices) are defined at input and limited by hardware only. The input is arranged assuming data is available for most years, but missing years are allowed. The model currently does not allow use of length data nor indices of survival rates. Diagnostics include index fits, residuals in catch and catch-at-age, and effective sample size calculations. Weights are input for different components of the objective function and allow for relatively simple age-structured production model type models up to fully parameterized models.

ASPM. Age-structured production models, also known as statistical catch-at-age (SCAA) models, are a technique of stock assessment that integrate fishery catch and fishery-independent sampling information. The procedures are flexible, allowing for uncertainty in the absolute magnitudes of catches as part of the estimation. Unlike virtual population analysis (VPA) that tracks the cumulative catches of various year classes as they age, ASPM is a forward projection simulation of the exploited population. ASPM is similar to the NOAA Fishery Toolbox applications ASAP (Age Structured Assessment Program) and SS2 (Stock Synthesis 2)

Availability. Refers to the distribution of fish of different ages or sizes relative to that taken in the fishery.

Biological reference points. Specific values for the variables that describe the state of a fishery system which are used to evaluate its status. Reference points are most often specified in terms of fishing mortality rate and/or spawning stock biomass. The reference points may indicate 1) a desired state of the fishery, such as a fishing mortality rate that will achieve a high level of sustainable yield, or 2) a state of the fishery that should be avoided, such as a high fishing mortality rate which risks a stock collapse and long-term loss of potential yield. The former type of reference points are referred to as "target reference points" and the latter are referred to as "limit reference points" or "thresholds". Some common examples of reference points are $F_{0.1}$, F_{MAX}, and F_{MSY}, which are defined later in this glossary.

B_0. Virgin stock biomass, i.e., the long-term average biomass value expected in the absence of fishing mortality.

B_{MSY}. Long-term average biomass that would be achieved if fishing at a constant fishing mortality rate equal to F_{MSY}.

Biomass Dynamics Model. A simple stock assessment model that tracks changes in stock using assumptions about growth and can be tuned to abundance data such as commercial catch rates, research survey trends or biomass estimates.

Catchability. Proportion of the stock removed by one unit of effective fishing effort (typically age-specific due to differences in selectivity and availability by age).

Control Rule. Describes a plan for pre-agreed management actions as a function of variables related to the status of the stock. For example, a control rule can specify how F or yield should vary with biomass. In the National Standard Guidelines (NSG), the "MSY control rule" is used to determine the limit fishing mortality, or Maximum Fishing Mortality Threshold (MFMT). Control rules are also known as "decision rules" or "harvest control laws."

Catch per Unit of Effort (CPUE). Measures the relative success of fishing operations, but also can be used as a proxy for relative abundance based on the assumption that CPUE is linearly related to stock size. The use of CPUE that has not been properly standardized for temporal-spatial changes in catchability should be avoided.

Exploitation pattern. The fishing mortality on each age (or group of adjacent ages) of a stock relative to the highest mortality on any age. The exploitation pattern is expressed as a series of values ranging from 0.0 to 1.0. The pattern is referred to as "flat-topped" when the values for all the oldest ages are about 1.0, and "dome-shaped" when the values for some intermediate ages are about 1.0 and those for the oldest ages are significantly lower. This pattern often varies by type of fishing gear, area, and seasonal distribution of fishing, and the growth and migration of the fish. The pattern can be changed by modifications to fishing gear, for example, increasing mesh or hook size, or by changing the proportion of harvest by gear type.

Mortality rates. Populations of animals decline exponentially. This means that the number of animals that die in an "instant" is at all times proportional to the number present. The decline is defined by survival curves such as: $N_{t+1} = N_t e^{-z}$

where N_t is the number of animals in the population at time t and N_{t+1} is the number present in the next time period; Z is the total instantaneous mortality rate which can be separated into deaths due to fishing (fishing mortality or F) and deaths due to all other causes (natural mortality or M) and e is the base of the natural logarithm (2.71828).To better understand the concept of an instantaneous mortality rate, consider the following example. Suppose the instantaneous total mortality rate is 2 (i.e., Z = 2) and we want to know how many animals out of an initial population of 1 million fish will be alive at the end of one year. If the year is apportioned into 365 days (that is, the 'instant' of time is one day), then 2/365 or 0.548% of the population will die each day. On the first day of the year, 5,480 fish will die (1,000,000 x 0.00548), leaving 994,520 alive. On day 2, another 5,450 fish die (994,520 x 0.00548) leaving 989,070 alive. At the end of the year, 134,593 fish $[1{,}000{,}000 \times (1 - 0.00548)^{365}]$ remain alive. If, we had instead selected a smaller 'instant' of time, say an hour, 0.0228% of the population would have died by the end of the first time interval (an hour), leaving 135,304 fish alive at the end of the year $[1{,}000{,}000 \times (1 - 0.00228)^{8760}]$. As the instant of time becomes shorter and shorter, the exact answer to the number of animals surviving is given by the survival curve mentioned above, or, in this example:

$N_{t+1} = 1{,}000{,}000e^{-2} = 135{,}335$ fish

Exploitation rate. The proportion of a population alive at the beginning of the year that is caught during the year. That is, if 1 million fish were alive on January 1 and 200,000 were caught during the year, the exploitation rate is 0.20 (200,000 / 1,000,000) or 20%.

F_{MAX}. The rate of fishing mortality that produces the maximum level of yield per recruit. This is the point beyond which growth overfishing begins.

$F_{0.1}$. The fishing mortality rate where the increase in yield per recruit for an increase in a unit of effort is only 10% of the yield per recruit produced by the first unit of effort on the unexploited stock (i.e., the slope of the yield-per-recruit curve for the $F_{0.1}$ rate is only one-tenth the slope of the curve at its origin).

$F_{10\%}$. The fishing mortality rate which reduces the spawning stock biomass per recruit (SSB/R) to 10% of the amount present in the absence of fishing. More generally, Fx%, is the fishing mortality rate that reduces the SSB/R to x% of the level that would exist in the absence of fishing.

F_{MSY}. The fishing mortality rate that produces the maximum sustainable yield.

Fishery Management Plan (FMP). Plan containing conservation and management measures for fishery resources, and other provisions required by the MSFCMA, developed by Fishery Management Councils or the Secretary of Commerce.

Generation Time. In the context of the National Standard Guidelines, generation time is a measure of the time required for a female to produce a reproductively-active female offspring for use in setting maximum allowable rebuilding time periods.

Growth overfishing. The situation existing when the rate of fishing mortality is above F_{MAX} and when fish are harvested before they reach their growth potential.

Limit Reference Points. Benchmarks used to indicate when harvests should be constrained substantially so that the stock remains within safe biological limits. The probability of exceeding limits should be low. In the National Standard Guidelines, limits are referred to as thresholds. In much of the international literature (e.g., FAO documents), "thresholds" are used as buffer points that signal when a limit is being approached.

Landings per Unit of Effort (LPUE). Analogous to CPUE and measures the relative success of fishing operations, but is also sometimes used a proxy for relative abundance based on the assumption that CPUE is linearly related to stock size.

MSFCMA. (Magnuson-Stevens Fishery Conservation and Management Act). U.S. Public Law 94-265, as amended through October 11, 1996. Available as NOAA Technical Memorandum NMFS-F/SPO-23, 1996.

Maximum Fishing Mortality Threshold (MFMT, $F_{THRESHOLD}$). One of the Status Determination Criteria (SDC) for determining if overfishing is occurring. It will usually be equivalent to the F corresponding to the MSY Control Rule. If current fishing mortality rates are above $F_{THRESHOLD,}$ overfishing is occurring.

Minimum Stock Size Threshold (MSST, $B_{THRESHOLD}$). Another of the Status Determination Criteria. The greater of (a) $½B_{MSY}$, or (b) the minimum stock size at which rebuilding to B_{MSY} will occur within 10 years of fishing at the MFMT. MSST should be measured in terms of spawning biomass or other appropriate measures of productive capacity. If current stock size is below $B_{THRESHOLD}$, the stock is overfished.

Maximum Spawning Potential (MSP). This type of reference point is used in some fishery management plans to define overfishing. The MSP is the spawning stock biomass per recruit (SSB/ R) when fishing mortality is zero. The degree to which fishing reduces the SSB/R is expressed as a percentage of the MSP (i.e., %MSP). A stock is considered overfished when the fishery reduces the %MSP below the level specified in the overfishing definition. The values of %MSP used to define overfishing can be derived from stock-recruitment data or chosen by analogy using available information on the level required to sustain the stock.

Maximum Sustainable Yield (MSY). The largest average catch that can be taken from a stock under existing environmental conditions.

Overfishing. According to the National Standard Guidelines, "overfishing occurs whenever a stock or stock complex is subjected to a rate or level of fishing mortality that jeopardizes the capacity of a stock or stock complex to produce MSY on a continuing basis." Overfishing is occurring if the MFMT is exceeded for 1 year or more.

Optimum Yield (OY). The amount of fish that will provide the greatest overall benefit to the Nation, particularly with respect to food production and recreational opportunities and taking into account the protection of marine ecosystems. MSY constitutes a "ceiling" for OY. OY may be lower than MSY, depending on relevant economic, social, or ecological factors. In the case of an overfished fishery, OY should provide for rebuilding to B_{MSY}.

Partial Recruitment. Patterns of relative vulnerability of fish of different sizes or ages due to the combined effects of selectivity and availability.

Rebuilding Plan. A plan that must be designed to recover stocks to the B_{MSY} level within 10 years when they are overfished (i.e. when B < MSST). Normally, the 10 years would refer to an expected time to rebuilding in a probabilistic sense.

Recruitment. This is the number of young fish that survive (from birth) to a specific age or grow to a specific size. The specific age or size at which recruitment is measured may correspond to when the young fish become vulnerable to capture in a fishery or when the number of fish in a cohort can be reliably estimated by a stock assessment.

Recruitment overfishing. The situation existing when the fishing mortality rate is so high as to cause a reduction in spawning stock which causes recruitment to become impaired.

Recruitment per spawning stock biomass (R/SSB). The number of fishery recruits (usually age 1 or 2) produced from a given weight of spawners, usually expressed as numbers of recruits per kilogram of mature fish in the stock. This ratio can be computed for each year class and is often used as an index of pre-recruit survival, since a high R/SSB ratio in one year indicates above-average numbers resulting from a given spawning biomass for a particular year class, and vice versa.

Reference Points. Values of parameters (e.g. B_{MSY}, F_{MSY}, $F_{0.1}$) that are useful benchmarks for guiding management decisions. Biological reference points are typically limits that should not be exceeded with significant probability (e.g., MSST) or targets for management (e.g., OY).

Risk. The probability of an event times the cost associated with the event (loss function). Sometimes "risk" is simply used to denote the probability of an undesirable result (e.g. the risk of biomass falling below MSST).

Status Determination Criteria (SDC). Objective and measurable criteria used to determine if a stock is being overfished or is in an overfished state according to the National Standard Guidelines.

Selectivity. Measures the relative vulnerability of different age (size) classes to the fishing gears(s).

Spawning Stock Biomass (SSB). The total weight of all sexually mature fish in a stock.

Spawning stock biomass per recruit (SSB/R or SBR). The expected lifetime contribution to the spawning stock biomass for each recruit. SSB/R is calculated assuming that F is constant over the life span of a year class. The calculated value is also dependent on the exploitation pattern and rates of growth and natural mortality, all of which are also assumed to be constant.

Stock Synthesis (SS). This application provides a statistical framework for calibration of a population dynamics model using a diversity of fishery and survey data. SS is designed to accommodate both age and size structure and with multiple stock sub-areas. Selectivity can be cast as age specific only, size-specific in the observations only, or size-specific with the ability to capture the major effect of size-specific survivorship. The overall model contains subcomponents which simulate the population dynamics of the stock and fisheries, derive the expected values for the various observed data, and quantify the magnitude of difference between observed and expected data. Parameters are searched for which will maximize the goodness-of-fit. A management layer is also included in the model allowing uncertainty in estimated parameters to be propagated to the management quantities, thus facilitating a description of the risk of various possible management scenarios. The structure of SS allows for building of simple to complex models depending upon the data available.

Survival Ratios. Ratios of recruits to spawners (or spawning biomass) in a stock-recruitment analysis. The same as the recruitment per spawning stock biomass (R/SSB), see above.

TAC. Total allowable catch is the total regulated catch from a stock in a given time period, usually a year.

Target Reference Points. Benchmarks used to guide management objectives for achieving a desirable outcome (e.g., OY). Target reference points should not be exceeded on average.

Uncertainty. Uncertainty results from a lack of perfect knowledge of many factors that affect stock assessments, estimation of reference points, and management. Rosenberg and Restrepo (1994) identify 5 types: measurement error (in observed quantities), process error (or natural population variability), model error (mis-specification of assumed values or model structure), estimation error (in population parameters or reference points, due to any of the preceding types of errors), and implementation error (or the inability to achieve targets exactly for whatever reason)

Virtual population analysis (VPA) (or cohort analysis). A retrospective analysis of the catches from a given year class which provides estimates of fishing mortality and stock size at each age over its life in the fishery. This technique is used extensively in fishery assessments.

Year class (or cohort). Fish born in a given year. For example, the 1987 year class of cod includes all cod born in 1987. This year class would be age 1 in 1988, age 2 in 1989, and so on.

Yield per recruit (Y/R or YPR). The average expected yield in weight from a single recruit. Y/R is calculated assuming that F is constant over the life span of a year class. The calculated value is also dependent on the exploitation pattern, rate of growth, and natural mortality rate, all of which are assumed to be constant.

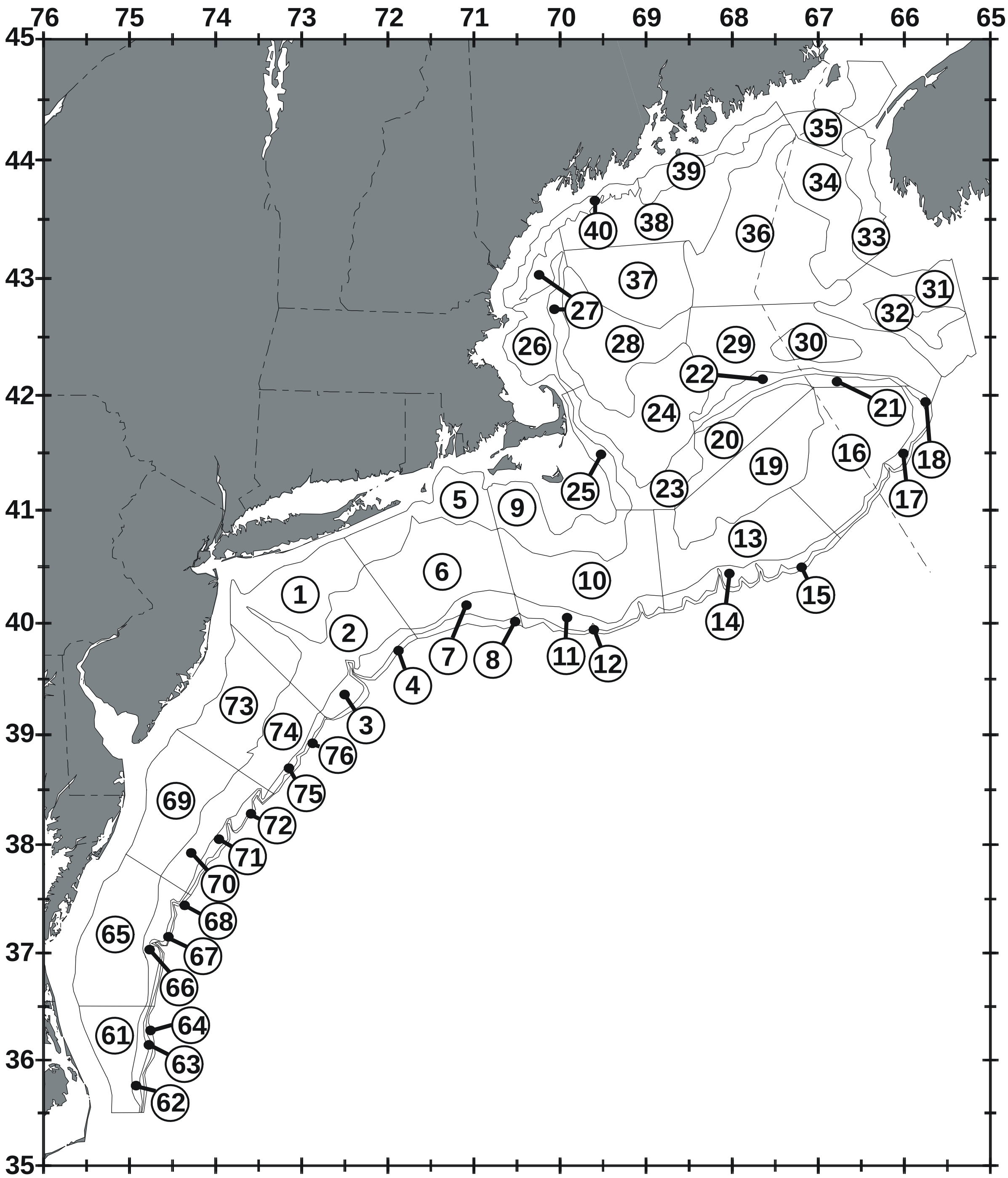


Figure 1. Offshore depth strata sampled during Northeast Fisheries Science Center bottom trawl research surveys. Some of these may not be sampled presently.

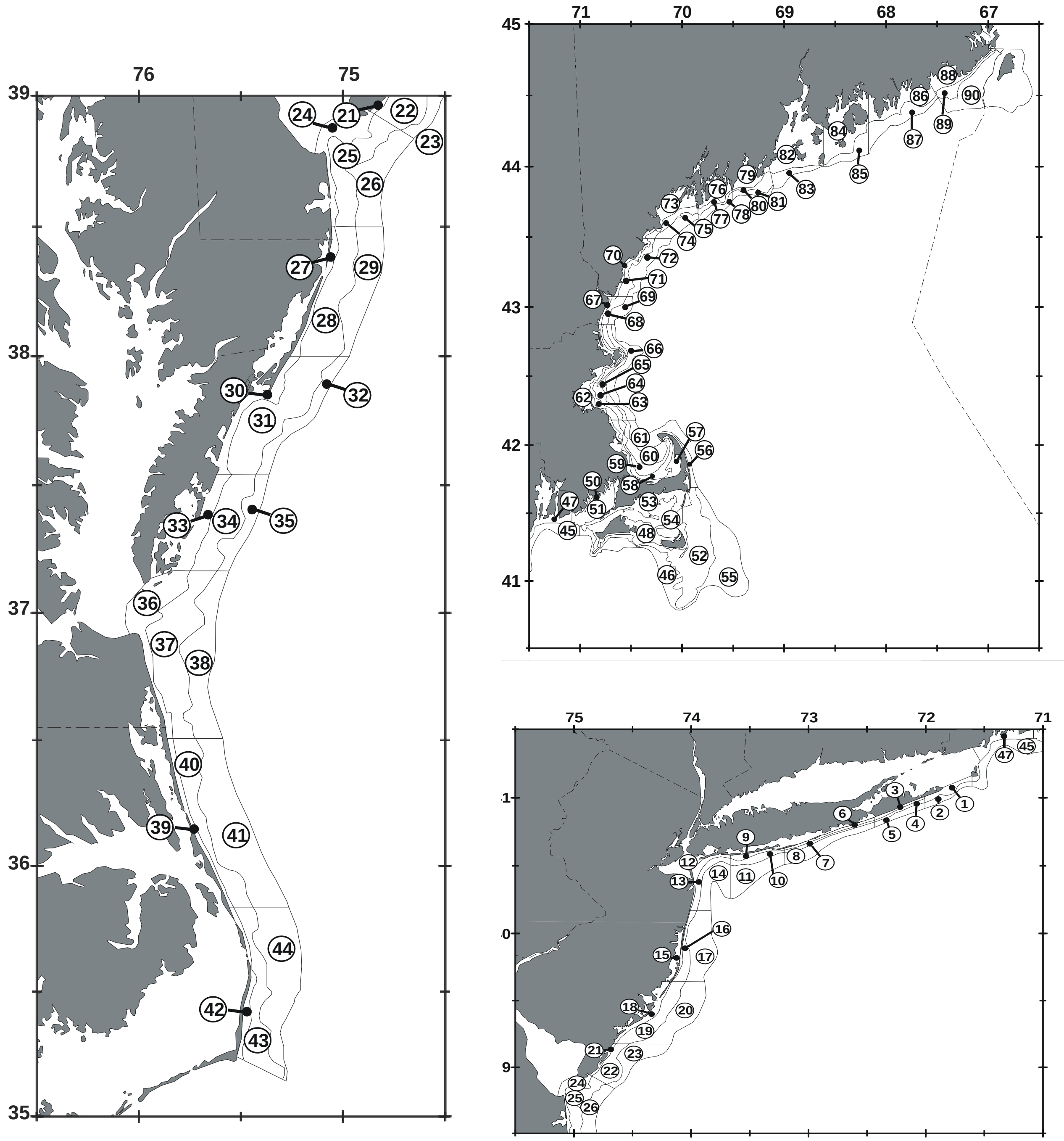
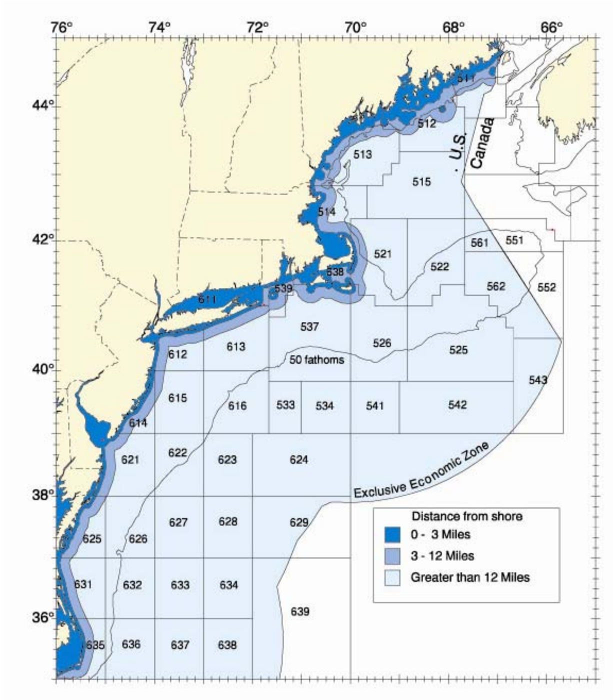


Figure 2. Inshore depth strata sampled during Northeast Fisheries Science Center bottom trawl research surveys. Some of these may not be sampled presently.



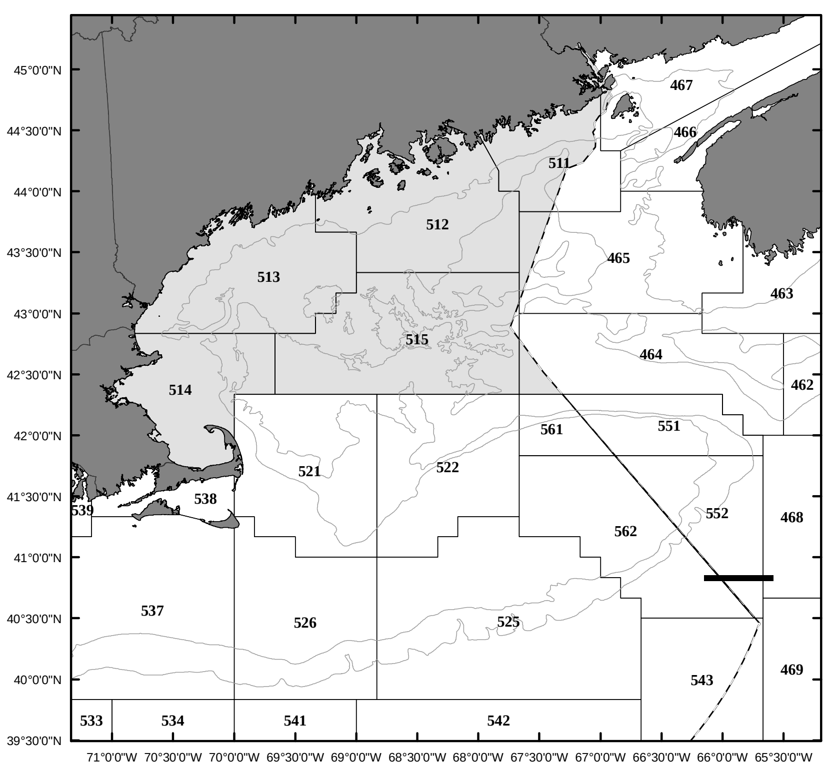


Figure 3. Statistical areas used for reporting commercial catches.

A. ATLANTIC HERRING ASSESSMENT SUMMARY FOR 2012

State of Stock: A statistical catch-at-age model, ASAP (Legault and Restrepo 1999), is proposed as the best scientific information available for determining the stock status for Atlantic herring. Spawning stock biomass (SSB) was estimated to be 517,930 mt in 2011 and fishing mortality rate at age 5 (F) was estimated to be 0.14 (Figure A1). Age 5 was used as the reference age for reporting fishing mortality rates because that age is fully selected in the mobile gear fleet, which accounted for most of the catches in recent years (see Catch and Status Table).

Maximum sustainable yield (MSY) reference points were estimated based on the fit of a Beverton-Holt stock-recruitment curve, which was estimated internally to ASAP. Steepness of the Beverton-Holt curve = 0.53, F_{MSY} = 0.27, SSB_{MSY} = 157,000 mt (½ SSB_{MSY} = 78,500), and MSY = 53,000 mt. Based on a comparison of the MSY reference points with the estimates of F and SSB for 2011, overfishing is not occurring and the stock is not overfished.

Projections: Short-term projections of future stock status were conducted based on the results of the ASAP model. The degree of retrospective error was sufficiently small, and did not warrant adjustment in the projections. Numbers-at-age in 2012 were drawn from 1000 vectors of numbers-at-age produced from MCMC simulations of the ASAP model. The projections assumed that catch in 2012 equaled the annual catch limit. Age-1 recruitment was based on the Beverton-Holt relationship estimated within ASAP. In general, results from several harvest scenarios suggested that overfishing will not occur and the stock will not become overfished through 2015. Results from the status quo catch projection were a notable exception because they resulted in small probabilities that overfishing could occur (Table A1).

Catch and Status Table: Atlantic herring

Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	Min[1]	Max[1]	Mean[1]
US Mobile Landings (000s mt)	93	102	94	93	103	81	84	103	67	81	67	124	99
US Fixed Landings (000s mt)	0.05	0.15	0.09	0.07	1.01	0.40	0.03	0.10	1.23	0.42	0.02	4.89	0.63
New Brunswick Weir Landings (000s mt)	12	9	21	13	13	31	6	4	11	4	4	31	15
Discards (000s mt)	0.04	0.03	0.49	0.30	0.20	0.06	0.53	0.46	0.26	0.17	0.03	0.55	0.25
Total Catch (000s mt)	105	111	115	107	117	112	91	108	79	85	79	145	115
Spawning Stock Biomass (000s mt)	433	371	371	410	376	367	385	301	313	518	301	840	468
Recruitment (millions age 1)	17,356	21,101	10,011	7,331	17,023	5,273	13,839	59,412	7,314	5,919	5,273	59,412	15,782
F	0.21	0.24	0.23	0.22	0.25	0.23	0.23	0.32	0.18	0.14	0.14	0.32	0.21

[1]Over the period 1996-2011, which is when natural mortality was increased.

Stock Distribution and Identification: The Gulf of Maine/Georges Bank Atlantic herring complex is composed of several spawning aggregations. Fisheries and research surveys, however, catch fish from a mix of the spawning aggregations and methods to distinguish fish from each aggregation are not yet well established. Consequently, recent assessments have combined data from all areas and conducted a single assessment of the entire complex. Although this approach poses a challenge to optimally managing each stock component and can create retrospective patterns within an assessment, the mixing of the spawning components in the fishery and surveys precludes separate assessments. Atlantic herring caught in the New Brunswick, Canada, weir fishery were considered part of the Gulf of Maine/Georges Bank complex because tagging studies suggested mixing. Herring from the Canadian Scotian Shelf stock also likely mix with the Gulf of Maine/Georges Bank complex, but the degree of mixing is unknown and methods to distinguish fish from each stock are not yet developed. Catches from the Scotian shelf were not considered part of the Gulf of Maine/Georges Bank complex. Despite a single assessment for the entire complex, catch limits are allocated to spatial management areas and catch allocations are based on estimates of stock composition and relative biomass among areas (Correia 2012).

Catches: US catch data were separated into two aggregate gear types, fixed and mobile gears, during 1964-2011. The reported catch is a sum of landings and self-reported discards, but discard estimates have only been available since 1996. Discards, however, were generally less than 1% of landings and do not represent a significant source of mortality (Wigley et al. 2011). Consequently, a lack of historical estimates of discards is not considered problematic to the assessment.

New Brunswick, Canada weir catches were provided for the years 1965-2011. Catches from this fishery were combined with US fixed gear catches for this assessment.

Catch in the US mobile gear fishery peaked in the late 1960s and early 70s, largely due to efforts from foreign fleets (Figure A2). Catch in this fishery has been relatively stable since about 2000 and has accounted for most of the Atlantic herring catches in recent years. Catch in the US fixed gear fishery has been variable, but has been relatively low since the mid-1980s (Figure A2). Catch in the NB weir fishery has also declined since the 1980s (Figure A2).

Total catches during 1964-2011 ranged from 44,613 mt in 1983 to 477,767 mt in 1968. Total catches during the past five years ranged from 79,413 mt in 2010 to 112,462 mt in 2007 and averaged 95,081 mt.

The US mobile gear fishery catches a relatively broad range of ages and some strong cohorts can be seen for several years. In contrast, the US fixed gear fishery and the NB weir fishery harvest almost exclusively age 2 herring.

Data and Assessment: The previous assessment of Atlantic herring used the statistical catch-at-age model ASAP and had a severe retrospective pattern (TRAC 2009). The new 2012 assessment also uses ASAP, but nearly all data inputs and model settings were reconsidered during development. Major changes to the input data are summarized here. Natural mortality during the 2009 TRAC was assumed to equal 0.2 for all ages and years. For this assessment, natural mortality was based on a combination of the Hoenig and Lorenzen methods, with the Hoenig method providing the scale of natural mortality and the Lorenzen method defining how natural mortality declined with age (Hoenig 1983; Lorenzen 1996). The natural mortality rates during 1996-2011 were increased by 50% to resolve a retrospective pattern and to ensure that the implied levels of consumption were consistent with observed increases in estimated consumption of herring. Consumption estimates were based on food habits data primarily for groundfish, but also informed by consumption estimates from marine mammals, highly migratory species, and seabirds. The 2009 TRAC also used catch data combined among all fishing gears and assumed selectivity equaled 1.0 for all ages. This assessment included separate catches and estimated selectivity separately for two aggregate gear types: fixed and mobile gears. This assessment also estimated selectivity for any survey with age composition data, which is in contrast to the 2009 TRAC which used age-specific indices. Finally, maturity at age varied through time in this assessment, but was constant among years in the 2009 TRAC. The time variation in maturity in this assessment was based on annual fits of general additive models to maturity data from males and females collected from commercial catches during July to September.

Abundances (i.e., arithmetic mean numbers per tow) from the NMFS spring, fall, and summer shrimp bottom trawl surveys were used in the assessment model along with annual coefficients of variation and age composition when they were available. The trawl door used on the spring and fall surveys changed in 1985 and likely altered the catchability of the survey gear. Consequently, the spring and fall surveys were split into two time series between 1984 and 1985, and these were treated as separate indices in assessment models. Calibrations were applied to the spring and fall surveys to account for changes in survey methods, including changes in research vessels.

Five other indices of abundance were considered, but not used in the final assessment model. These indices included: NMFS winter survey, NMFS herring acoustic survey, Massachusetts state surveys (spring and fall), joint Maine/New Hampshire state surveys (spring and fall), and a larval index of abundance.

Biological Reference Points (BRPs): Updated MSY reference points were estimated based on the fit to a Beverton-Holt stock-recruitment curve, which was estimated internally to the ASAP model. Steepness of the Beverton-Holt curve = 0.53. For calculating MSY reference points, ASAP used the inputs (e.g., weights at age, M) from the terminal year of the assessment (i.e., 2011). Using inputs from the terminal year of the assessment had the consequence of using natural mortality rates from the period when these rates were increased by 50% (see Data and Assessment). Estimates of MSY BRPs were: F_{MSY} = 0.27, SSB_{MSY} = 157,000 mt (½ SSB_{MSY} = 78,500), and MSY = 53,000 mt.

MSY reference points from the previous assessment (TRAC 2009) were based on the fit of a Fox surplus production model (TRAC 2009), and F_{MSY} = 0.27, SSB_{MSY} = 670,600 mt (½ SSB_{MSY} = 335,300 mt) and MSY = 178,000 mt.

BRPs changed since the previous assessment primarily because the Fox model had been used during the 2009 TRAC and assumed natural mortality rates were revised.

Fishing Mortality: F at age-5 equaled 0.14 in 2011, and was near the all-time low of 0.13 (1994) (Figure A3). F in 2011, however, was not representative of fishing mortality rates in recent years, which averaged 0.23 during 2000-2009 and also showed an increasing trend during those years (Figure A3). Fishing mortality rates in 2010 and 2011 were relatively low due to the presence of a strong cohort which increased the stock biomass (see below). The maximum F over the time series was 0.80 in 1980 (Figure A3).

Biomass: Based on the ASAP model, SSB = 517,930 mt in 2011. Over the entire time series, SSB ranged 53,349 mt in 1978 to 839,710 mt in 1997 (Figure A4). SSB declined during 1997-

2010, but increased in 2011 (Figure A4). Estimated total January 1 biomass was 1,322,446 mt in 2011, and ranged from a minimum of 180,527 mt in 1982 to a maximum of 1,936,769 mt in 2009 (Figure A4). Total biomass and SSB showed similar trends over time, but with 1-2 year lags caused by total biomass being more reflective of immature recruits than SSB. Spawning stock and total biomass increased after 2009, mostly due to the presence of a strong cohort (see below).

Recruitment: With the exception of 2009, age-1 recruitment since 2006 has been below the 1996-2011 average of 15.8 billion fish (Figure A5). The 2009 age-1 recruitment, however, was the largest in the time series at 59.4 billion fish. This large 2009 age-1 cohort consistently appeared in all sources of data that contain age composition.

Special Comments:

- This assessment represents a significant change from previous assessments. Unlike previous assessments, the catch at age was partitioned into mobile and fixed gear fleets and treated separately in a new formulation of the ASAP model. Age-specific and time-varying natural mortality rates were developed. Estimates of herring consumption by a representative suite of predators justified a 50% increase in natural mortality beginning in 1996, which implies a decrease in sustainable yield.
- The assessment was evaluated for uncertainty and robustness to various parameters. The justification for the 50% increase in natural mortality (M) beginning in 1996 was further evaluated using alternative increases of 0%, 30%, 40%, 60%, 70%, and a reduction in the average M among ages in each year from 0.3 (as in the base model) to 0.2. Based on fits to data, degree of retrospective pattern, and general similarity between levels of implied consumption to estimates of consumption, the 50% increase used in the base model was considered appropriate.
- The steepness parameter of the stock-recruitment model was also profiled across a range of values. This profile suggested that the data did not contain much information about the appropriate value for steepness and that subsequent biological reference points were also highly uncertain. For example, over approximate 95% confidence intervals for steepness (0.35-0.85), MSY ranged from 40,000 to 78,000 mt, SSB_{msy} ranged from 73,000 to 277,000 mt, and F_{msy} ranged from 0.12 to 0.7. Stock status in 2011, however, was robust to this uncertainty, with a broad range of comparisons resulting in the conclusion of not overfished and no overfishing ($SSB > ½ SSB_{MSY}$ and $F < F_{msy}$). Only in the extreme case of steepness equal to 0.35, which was considered implausible, would overfishing be occurring. Similarly, sensitivity runs of projections through 2015 based inputs and results of the current assessment, mostly over a range of assumptions about natural mortality, suggested that the probability of the stock being overfished or for overfishing to occur using commonly applied harvest scenarios (e.g., F_{MSY}, MSY) was generally zero.
- The robust nature of stock status was likely driven by the age-1 cohort in 2009, which was estimated to be the largest on record. To test the sensitivity of stock status to the presence of this cohort, projections through 2015 at F_{MSY} were conducted with the size of that cohort cut in half, which made the age-1 2009 cohort approximately equal to previous high recruitments. The probability of the stock being overfished or for overfishing to occur remained at zero. Furthermore, a sensitivity run was conducted with

the variation of the annual recruitments from the underlying Beverton-Holt relationship being more restricted than in the base model (CV in base =1, CV in sensitivity = 0.67). This sensitivity suggested that even with these additional restrictions on recruitment variation, the age-1 2009 cohort would still be the largest on record.

- Natural mortality is an uncertainty in this assessment. Of particular importance is acceptance of the scale of the herring consumption estimates (Figure A6). The 50% increase in natural mortality from the original natural mortality values during 1996-2011 used in the ASAP model was employed to reduce retrospective patterns in SSB and to make implied biomass removals from input natural mortality rates and the consumption data more consistent.
- The reference points and projections were based on the assumption that prevailing conditions would persist, including the relatively high natural mortality rates of 1996-2011. If life history traits such as natural mortality change rapidly, and prevailing conditions become altered, the associated biological reference points and projections would likewise need to be reexamined.
- In the short-term, the 2009 age-1 cohort (2008 year class) may reduce the vulnerability of this stock to overfishing. The strength of large cohorts, however, is often overestimated in the short-term. Consequently, the strength of this cohort should be interpreted cautiously and decisions based on this assessment should consider this uncertainty.
- Recent annual catches have been well above MSY. Consistent with this observation, SSB has declined since 1996 with the exception of recent increases driven by the 2009 age-1 cohort. The reference points (e.g., MSY), however, are uncertain.

References:

Correia, S. 2012. Simulating removals of the inshore component for the herring fishery for years 1999-2008 using the herring PDT's risk assessment methodology. Appendix II of the 2010-2012 Atlantic Herring Fishery Specifications. New England Fishery Management Council.

Hoenig, J.M. 1983. Empirical use of longevity data to estimate mortality rates. Fishery Bulletin 82(1): 898-903.

Legault, C. M., and V. R. Restrepo. 1999. A flexible forward age-structured assessment program. ICCAT Coll. Vol. Sci. Pap. 49(2): 246-253.

Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology 49: 627-647.

TRAC (Transboundary Resources Assessment Committee) 2009. Transboundary Resources Assessment Committee Gulf of Maine/Georges Bank Atlantic herring stock assessment update. TRAC reference document 2009/04.

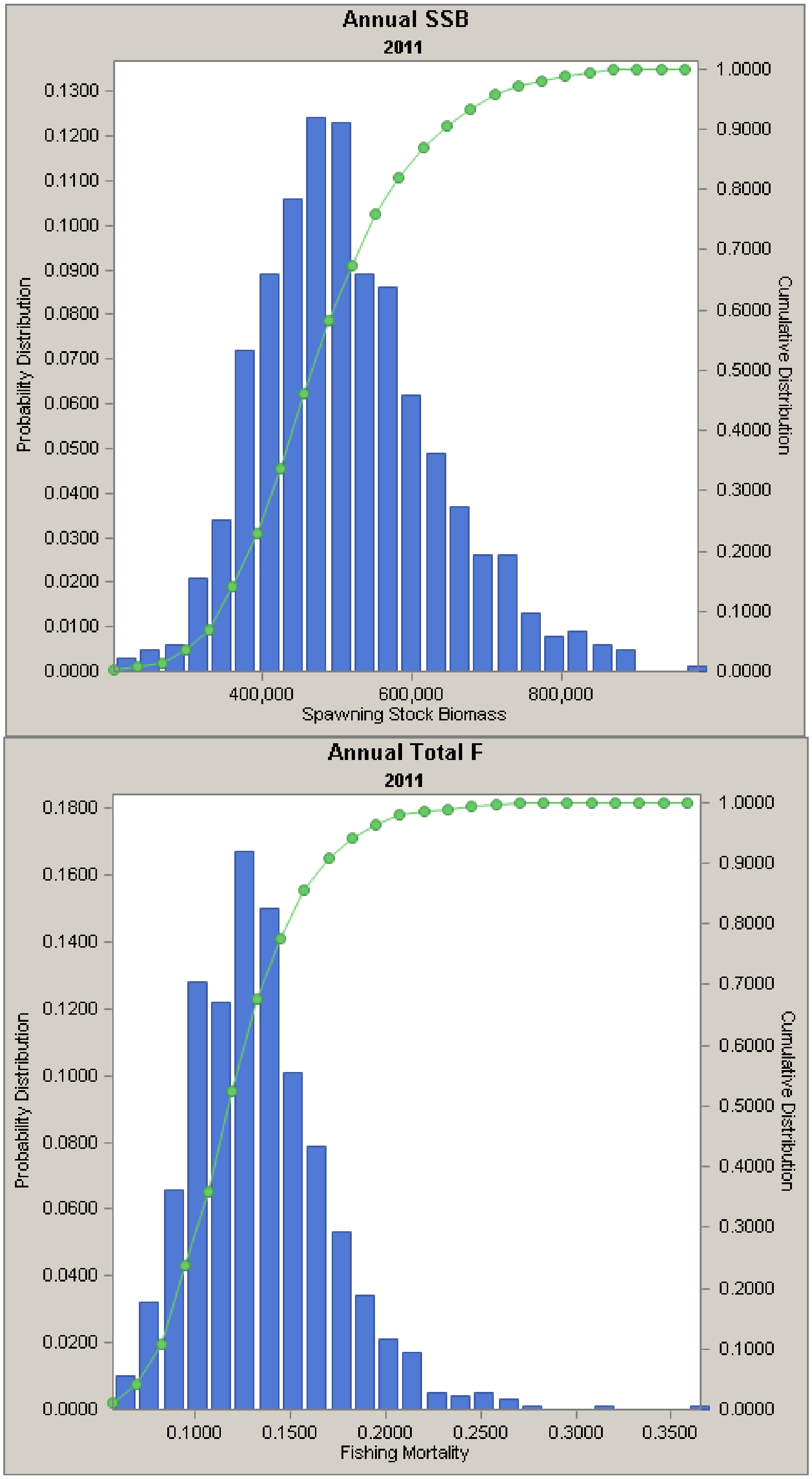
Wigley, S.E., J. Blaylock, P.J. Rago, J. Tang, H.L. Haas, and G. Shield. 2011. Standardized bycatch reporting methodology 3-year review report – 2011 Part 1. Northeast Fisheries Science Center Reference Document 11-09.

A. Atlantic Herring – Tables

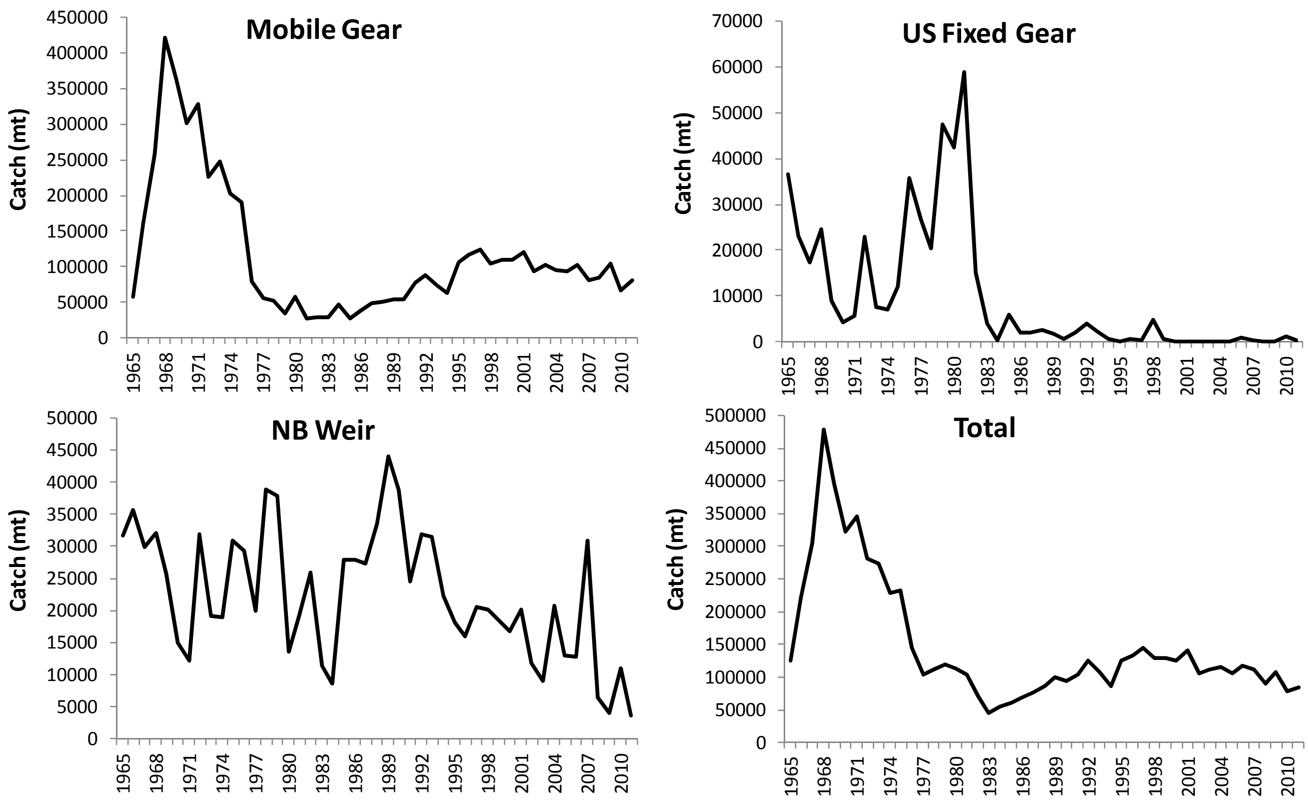
Table A1. Results of three-year Atlantic herring projections for the base ASAP model.

Fmsy = 0.267	SSBmsy = 157,000 mt	steepness = 0.53	MSY = 53,000 mt
2011 F (age 5)	SSB 2011		2011 catch
0.14	518,000 mt		85,000 mt
2012 catch = 87,683 mt (quota)			
	2013	**2014**	**2015**
	F_{msy}		
F	0.267	0.267	0.267
SSB	496,064 mt	368,501 mt	308,949 mt
80% CI	362,965 - 688,585 mt	275,695 - 517-815 mt	237,755 - 411,808 mt
Prob < SSBmsy/2	0	0	0
catch	168,775 mt	126,589 mt	104,430 mt
80% CI	124,868 - 230,764 mt	95,835 - 171,145 mt	79,505 - 139,925 mt
	$F_{75\%\ msy}$		
F	0.2	0.2	0.2
SSB	523,243 mt	409,309 mt	354,559 mt
80% CI	382,573 - 723,975 mt	306,011 - 574,128 mt	272,751 - 473,021 mt
Prob < SSBmsy/2	0	0	0
catch	130,025 mt	102,470 mt	87,574 mt
80% CI	96,216 - 177,894 mt	77,476 - 138,665 mt	66,739 - 117,318 mt
	$F_{status\ quo}$		
F	0.14	0.14	0.14
SSB	548,788 mt	450,496 mt	402,551 mt
80% CI	401,571 - 760,028 mt	336,594 - 631,502 mt	309,334 - 537,414 mt
Prob < SSBmsy/2	0	0	0
catch	93,159 mt	76,823 mt	67,912 mt
80% CI	68,954 - 127,518 mt	58,022 - 104,055 mt	51,752 - 91,001 mt
	MSY		
F	0.08	0.09	0.1
80% CI	0.06 - 0.11	0.07 - 0.12	0.07 - 0.14
Prob > Fmsy	0	0	0
SSB	576,092 mt	492,162 mt	448,725 mt
80% CI	413,046 - 813,298 mt	351,530 - 716,931 mt	321,209 - 633,132 mt
Prob < SSBmsy/2	0	0	0
catch	53,000 mt	53,000 mt	53,000 mt
	Status quo catch		
F	0.13	0.16	0.19
80% CI	0.1 - 0.18	0.11 - 0.23	0.13 - 0.27
Prob > Fmsy	1%	4%	10%
SSB	551,686 mt	446,496 mt	385,995 mt
80% CI	388,989 - 789,568 mt	306,349 - 669,721 mt	259,178 - 569,560 mt
Prob < SSBmsy/2	0	0	0
2012 quota	87,683 mt	87,683 mt	87,683 mt

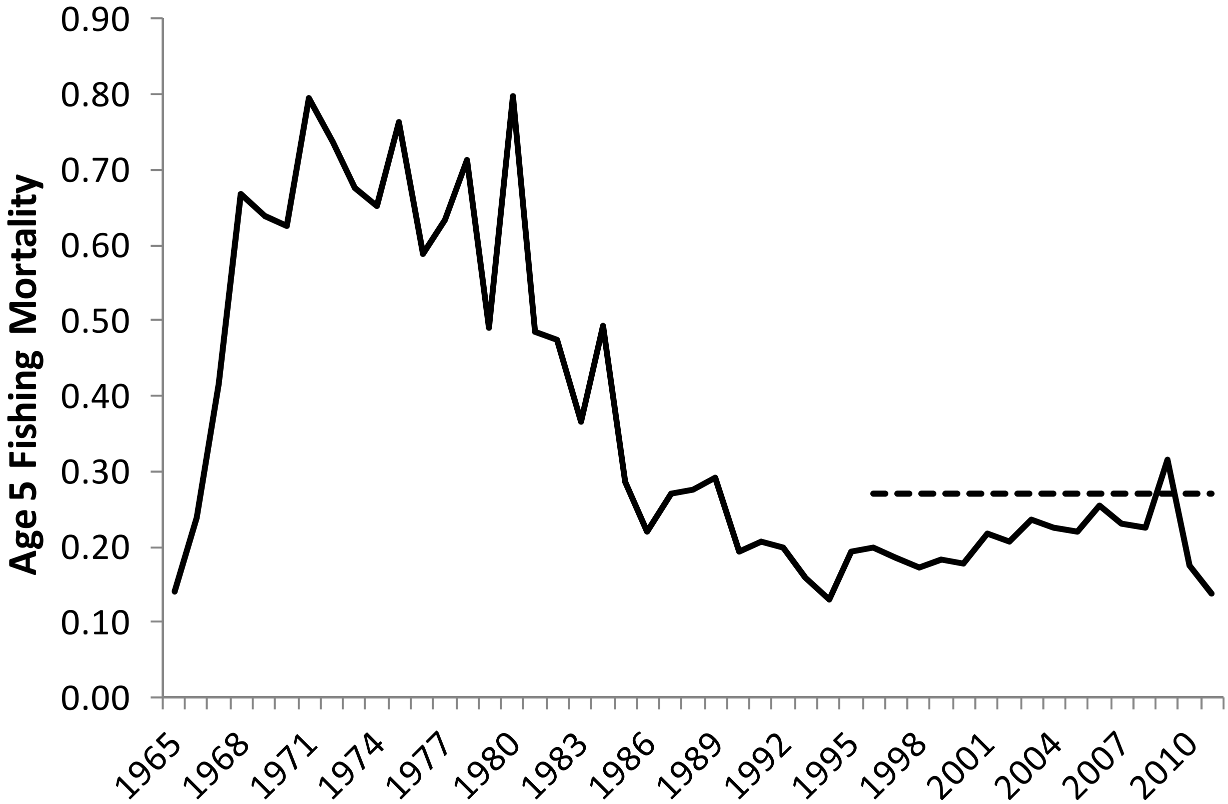
A. Atlantic Herring – Figures



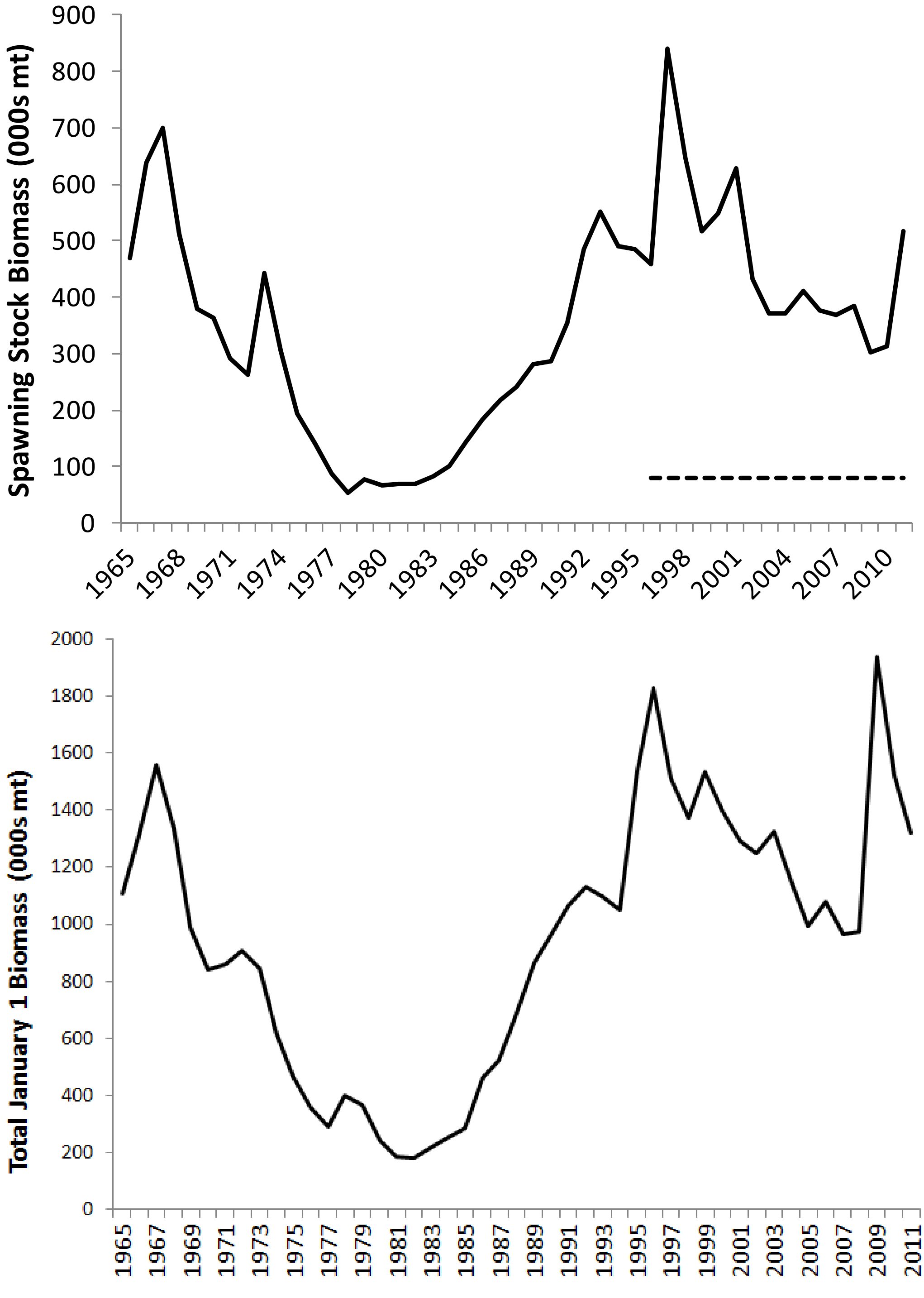
A1. Posterior densities of Atlantic herring SSB and F in 2011 from the ASAP base run.



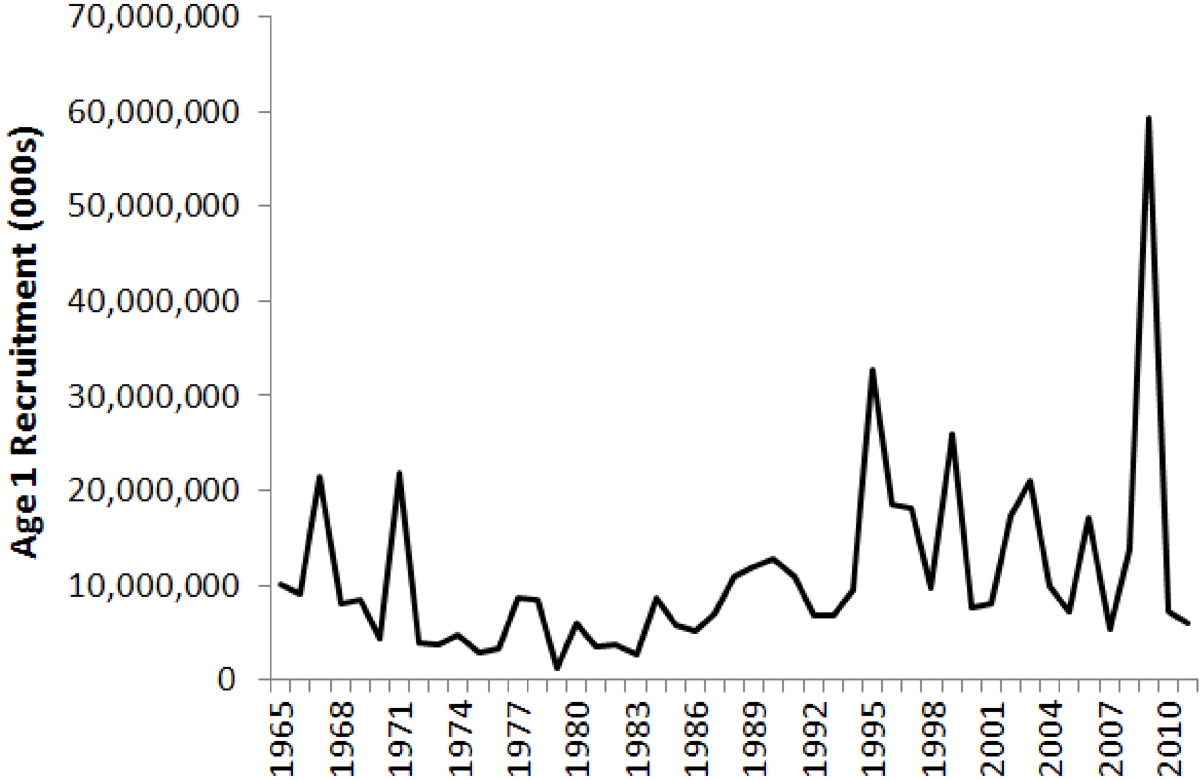
A2. Atlantic herring catch (mt) during 1965-2011 for US mobile gears, US fixed gears, NB weir fishery, and total catch. Discards estimates were only available since 1996.



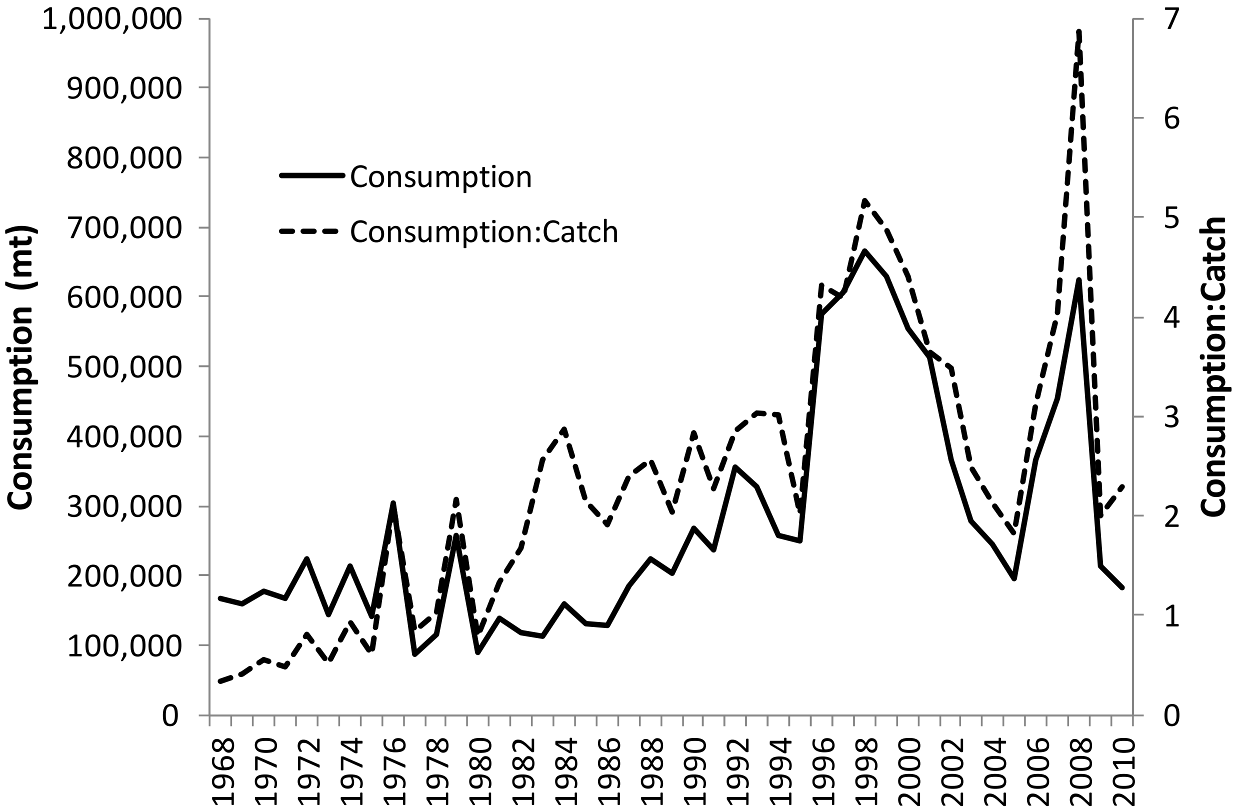
A3. Atlantic herring age-5 fishing mortality (solid line) and F_{MSY} (dashed line) estimated from the ASAP model base run. The F_{MSY} reference line is only provided during 1996-2011 because the reference point from this assessment is only for this time period.



A4. Atlantic herring spawning stock biomass (000s MT; solid line; top panel), ½ SSB_{MSY} (dashed line; top panel), and total biomass (000s MT; bottom panel) time series estimated from the ASAP base run. The ½ SSB_{MSY} reference line is shown for 1996-2011 because the reference point from this assessment is only for this time period.



A5. Atlantic herring age-1 recruitment (000s) over time, estimated from the ASAP model base run.



A6. Consumption of Atlantic herring by groundfish species, marine mammals, highly migratory species and seabirds (solid line). Also shown, the ratio of consumption to fishery catch (dashed line), 1968-2010.

Appendix: Stock Assessment Terms of Reference for SAW/SARC54 (June 5-9, 2012)

(file vers.: 10/21/11b)

A. Atlantic herring

1. Estimate catch from all sources including landings and discards. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.

2. Present the survey data being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, larval surveys, age-length data, predator consumption rates, etc.). Investigate the utility of commercial LPUE as a measure of relative abundance, and characterize the uncertainty and any bias in these sources of data.

3. Evaluate the utility of the NEFSC fall acoustic survey to the stock assessment of herring. Consider degree of spatial and temporal overlap between the survey and the stock. Compare acoustic survey results with measures derived from bottom trawl surveys.

4. Evaluate the validity of the current stock definition, and determine whether it should be changed. Take into account what is known about migration among stock areas.

5. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-6), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

6. Consider the implications of consumption of herring, at various life stages, for use in estimating herring natural mortality rate (M) and to inform the herring stock-recruitment relationship. Characterize the uncertainty of the consumption estimates. If possible integrate the results into the stock assessment.

7. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

8. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model, should one be developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).
    a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.
    b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-7).

9. Using simulation/estimation methods, evaluate consequences of alternative harvest policies in light of uncertainties in model formulation, presence of retrospective patterns, and incomplete information on magnitude and variability in M.

10. Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).
    a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).
    b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
    c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

11. For any research recommendations listed in recent peer reviewed assessment and review panel reports, review, evaluate and report on the status of those research recommendations. Identify new research recommendations.

Appendix to the SAW Assessment TORs:

Clarification of Terms used in the SAW/SARC Terms of Reference

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidel. Fed. Reg., v. 74, no. 11, 1-16-2009):

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of [overfishing limit] OFL and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, [optimal yield] OY does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

"Vulnerability. A stock's vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce MSY and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality)." (p. 3205)

Rules of Engagement among members of a SAW Assessment Working Group:

Anyone participating in SAW assessment working group meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

Procedures for Issuing Manuscripts
in the
***Northeast Fisheries Science Center Reference Document* (CRD) Series**

Clearance

All manuscripts submitted for issuance as CRDs must have cleared the NEFSC's manuscript/abstract/webpage review process. If any author is not a federal employee, he/she will be required to sign an "NEFSC Release-of-Copyright Form." If your manuscript includes material from another work which has been copyrighted, then you will need to work with the NEFSC's Editorial Office to arrange for permission to use that material by securing release signatures on the "NEFSC Use-of-Copyrighted-Work Permission Form."

For more information, NEFSC authors should see the NEFSC's online publication policy manual, "Manuscript/abstract/webpage preparation, review, and dissemination: NEFSC author's guide to policy, process, and procedure," located in the Publications/Manuscript Review section of the NEFSC intranet page.

Organization

Manuscripts must have an abstract and table of contents, and (if applicable) lists of figures and tables. As much as possible, use traditional scientific manuscript organization for sections: "Introduction," "Study Area" and/or "Experimental Apparatus," "Methods," "Results," "Discussion," "Conclusions," "Acknowledgments," and "Literature/References Cited."

Style

The CRD series is obligated to conform with the style contained in the current edition of the United States Government Printing Office Style Manual. That style manual is silent on many aspects of scientific manuscripts. The CRD series relies more on the CSE Style Manual. Manuscripts should be prepared to conform with these style manuals.

The CRD series uses the American Fisheries Society's guides to names of fishes, mollusks, and decapod crustaceans, the Society for Marine Mammalogy's guide to names of marine mammals, the Biosciences Information Service's guide to serial title abbreviations, and the ISO's (International Standardization Organization) guide to statistical terms.

For in-text citation, use the name-date system. A special effort should be made to ensure that all necessary bibliographic information is included in the list of cited works. Personal communications must include date, full name, and full mailing address of the contact.

Preparation

Once your document has cleared the review process, the Editorial Office will contact you with publication needs – for example, revised text (if necessary) and separate digital figures and tables if they are embedded in the document. Materials may be submitted to the Editorial Office as files on zip disks or CDs, email attachments, or intranet downloads. Text files should be in Microsoft Word, tables may be in Word or Excel, and graphics files may be in a variety of formats (JPG, GIF, Excel, PowerPoint, etc.).

Production and Distribution

The Editorial Office will perform a copy-edit of the document and may request further revisions. The Editorial Office will develop the inside and outside front covers, the inside and outside back covers, and the title and bibliographic control pages of the document.

Once both the PDF (print) and Web versions of the CRD are ready, the Editorial Office will contact you to review both versions and submit corrections or changes before the document is posted online.

A number of organizations and individuals in the Northeast Region will be notified by e-mail of the availability of the document online.

Research Communications Branch
Northeast Fisheries Science Center
National Marine Fisheries Service, NOAA
166 Water St.
Woods Hole, MA 02543-1026

MEDIA
MAIL

Publications and Reports of the Northeast Fisheries Science Center

The mission of NOAA's National Marine Fisheries Service (NMFS) is "stewardship of living marine resources for the benefit of the nation through their science-based conservation and management and promotion of the health of their environment." As the research arm of the NMFS's Northeast Region, the Northeast Fisheries Science Center (NEFSC) supports the NMFS mission by "conducting ecosystem-based research and assessments of living marine resources, with a focus on the Northeast Shelf, to promote the recovery and long-term sustainability of these resources and to generate social and economic opportunities and benefits from their use." Results of NEFSC research are largely reported in primary scientific media (*e.g.*, anonymously-peer-reviewed scientific journals). However, to assist itself in providing data, information, and advice to its constituents, the NEFSC occasionally releases its results in its own media. Currently, there are three such media:

NOAA Technical Memorandum NMFS-NE -- This series is issued irregularly. The series typically includes: data reports of long-term field or lab studies of important species or habitats; synthesis reports for important species or habitats; annual reports of overall assessment or monitoring programs; manuals describing program-wide surveying or experimental techniques; literature surveys of important species or habitat topics; proceedings and collected papers of scientific meetings; and indexed and/or annotated bibliographies. All issues receive internal scientific review and most issues receive technical and copy editing.

Northeast Fisheries Science Center Reference Document -- This series is issued irregularly. The series typically includes: data reports on field and lab studies; progress reports on experiments, monitoring, and assessments; background papers for, collected abstracts of, and/or summary reports of scientific meetings; and simple bibliographies. Issues receive internal scientific review and most issues receive copy editing.

Resource Survey Report (formerly *Fishermen's Report*) -- This information report is a regularly-issued, quick-turnaround report on the distribution and relative abundance of selected living marine resources as derived from each of the NEFSC's periodic research vessel surveys of the Northeast's continental shelf. This report undergoes internal review, but receives no technical or copy editing.

TO OBTAIN A COPY of a *NOAA Technical Memorandum NMFS-NE* or a *Northeast Fisheries Science Center Reference Document*, either contact the NEFSC Editorial Office (166 Water St., Woods Hole, MA 02543-1026; 508-495-2350) or consult the NEFSC webpage on "Reports and Publications" (http://www.nefsc.noaa.gov/nefsc/publications/). To access *Resource Survey Report*, consult the Ecosystem Surveys Branch webpage (http://www.nefsc.noaa.gov/femad/ecosurvey/mainpage/).

54th Northeast Regional Stock Assessment Workshop (SAW 54)

Stock Assessment Review Committee (SARC) Meeting
5 - 9 June 2012
Northeast Fisheries Science Center
Wood's Hole, Mass

SARC 54 PANEL SUMMARY REPORT

27 June 2012

Review Panel
R. O'Boyle (chair)
C. Francis
N. Hall
N. Klaer

TABLE OF CONTENTS

INTRODUCTION

Background

The Stock Assessment Review Committee (SARC) of the 54th Stock Assessment Workshop (SAW 54) met at the Northeast Fisheries Science Center (NEFSC), Woods Hole, MA during 5th – 9th June 2012 to review Northeast regional benchmark stock assessments of Atlantic Herring *(Clupea harengus)* and Southern New England / Mid-Atlantic Bight Yellowtail Flounder (*Limanda ferruginea*), guided by the SAW 54 Terms of Reference (Annex 2 of the SAW 54 Statement of Work provided below).

The SARC review panel (herein called the "Panel") consisted of R. O'Boyle (Beta Scientific Consulting, Canada) and three scientists affiliated with the Center of Independent Experts: C. Francis, N. Hall and N. Klaer. R. O'Boyle is also a member of the New England Fisheries Management Council's Scientific and Statistical Committee (SSC).

The SARC was assisted by the NEFSC SAW Chairman, J. Weinberg, A. O'Brien, and P. Rago (NEFSC). Documentation for the herring assessment was prepared by the NEFSC Herring Working Group (HWG), and the presentations at the meeting were made by J. Deroba (NEFSC). Documentation for the yellowtail assessment was prepared by the NEFSC Southern Demersal Working Group (SDWG), and the presentation at the meeting was made by L. Alade (NEFSC).

The rapporteurs who recorded the discussion to assist the Panel in its deliberations were T. Chute (herring) and J. Blaylock (yellowtail).

Review of Activities

The Northeast Regional Stock Assessment Workshop (SAW) is a process consisting of three phases:

1. preparation of stock assessments by SAW Working Groups and/or by ASMFC Technical Committees / Assessment Committees
2. SARC peer review of assessments by a panel of external experts who judge the adequacy of the assessments as a basis of scientific advice to managers
3. presentation of the results and reports to the Northeast Region's fishery management bodies

Regarding the first phase of the SAW process, the previous herring assessment was conducted during 8 – 11 June 2009 by the Transboundary Resource Assessment Committee (TRAC), which in turn used the model formulation developed at a 2006 TRAC benchmark review. To prepare the herring assessment, the HWG held a data review meeting (30 Jan – 3 Feb 2012) and a models meeting (9 – 13 April 2012). The previous SNE/MAB yellowtail assessment benchmark was conducted during the November 2007 – August 2008 GARM III review. To prepare the yellowtail assessment, the SDWG held an industry meeting (27 February 2012), a data review meeting (2 – 4 April 2012) and a models meeting (30 April – 4 May 2012).

Regarding the second phase of the SAW process, in May 2012, the NEFSC provided the Panel with background documentation on each species. The first assessment reports from the working groups were made available to the Panel on 18 May 2012 with a complete set of working papers available soon thereafter (see bibliography below).

The SARC was convened during 5th – 9th June 2012, the agenda of which is summarized in Table 1 (see full agenda in annex 3 of the SAW Statement of Work below).

Table 1. Summary of SARC/SAW 54 Agenda during 5th – 9th June 2012

<table>
<tr><th></th><th>5th June</th><th>6th June</th><th>7th June</th><th>8th June</th><th>9th June</th></tr>
<tr><th></th><th>Tuesday</th><th>Wednesday</th><th>Thursday</th><th>Friday</th><th>Saturday</th></tr>
<tr><td>Morning</td><td>Panel meeting (10:30)</td><td>Herring assessment</td><td>Assessment revisits</td><td>Summary Report Review</td><td>Panel Report Discussion (closed meeting)</td></tr>
<tr><td rowspan="2">Afternoon</td><td rowspan="2">Herring Assessment</td><td rowspan="2">SNE Yellowtail assesment</td><td>Assessment revisits</td><td>Summary Report Review</td><td rowspan="2"></td></tr>
<tr><td>Summary Report Review</td><td>Panel Report Discussion (closed meeting)</td></tr>
</table>

The Panel devoted Friday afternoon and Saturday morning to developing consensus points for each stocks' Terms of Reference as well as observations on the SARC process. It was agreed that each panelist would use these points to draft a section of this report, which would then be compiled and edited by the SARC chair. The report was distributed to the Panel for final review before being submitted to the NEFSC.

SARC Process

The Panel reached consensus on all Terms of Reference for each stock. It acknowledges the significant work that the two assessment working groups had undertaken to prepare for the SARC review. It also appreciates the professionalism and cooperation of NEFSC staff at the SARC meeting which significantly assisted the peer review. Notwithstanding this, during the course of the review, issues came the attention of the Panel which were not specific to the assessment of either species, resolution of which would assist future SARC reviews. These relate to both the terms of reference of reviews and the presentation of assessment results.

The terms of reference of the herring and yellowtail assessments required a review of their stock definitions. These were conducted during the data meetings of each species, the results of which were brought forward to the assessment meeting. Changes in stock definition have consequences throughout the management system and should not be undertaken without significant consideration of all sources of information. One could expect that there would be considerable reluctance to change stock definition without substantial evidence to the contrary. The Panel considers that reviews of stock definition would more productively be undertaken outside of the normal assessment process and on a schedule which would allow significant changes if these were felt warranted. The review of stock definition needs to highlight the uncertainties in the interactions amongst populations that might influence the interpretation of data during an assessment. This review also needs to determine the catch and indices appropriate for the stock(s) in question.

The term of reference for each stock included review of the data (e.g. catch, indices) to be used in the assessments. Consideration should be given to formally separating the data review from that of the assessment, similar to the SEDAR process conducted by the NMFS Southeast Fisheries Science Center. This would allow the peer review of the assessment to fully devote its attention to determination of stock status, reference points and projections. The data reviews could be undertaken on groups of species (e.g. groundfish, pelagics) to obtain data perspectives across stocks. These reviews would indicate the relative reliability of various datasets (e.g. survey indices) and clearly specify which data and their uncertainties should be brought forward to the assessment review.

While it was evident that the HWG and SDWG had spent considerable time preparing the documents and presentations for the SARC 54, there was unevenness in the relative content of each. In one stock, information was summarized in the working papers that were not summarized in the presentations at the SARC review while in the other stock, the reverse was the case. Greater detail was sometimes available in the presentation than was in the working papers. It would assist future peer reviews if the evidence supporting the conclusions of a working group were both presented in the assessment report and the presentation such that panel reviewers may assess whether or not the conclusions are justified.

On some of the figures describing current stock status, both the posterior distribution of the relevant indicator (e.g. spawning stock biomass), as determined through an MCMC process, and point estimates were displayed. Where MCMC was used to provide distributions of current stock status, these should be used to provide metrics and their uncertainty required by the management system. In this case, coefficients of variations based on the Hessian matrix are not required to be reported.

In both assessments, historical trends in spawning stock biomass and fishing mortality were illustrated along with current estimates of the biological reference points (BRPs). However, as highlighted by both assessments, temporal changes in population processes can dramatically change the BRPs. Changes in growth and fishery selectivity can also change the BRPs. There are a number of ways temporal changes in BRPs can be displayed (annually, smoothed over a number of years, by decade, etc). It would be useful for the NEFSC to develop a policy for the estimation and presentation of temporal changes in BRPs to both avoid future confusion and promote transparent communication with stakeholders and managers on long-term productivity and fisheries changes.

ATLANTIC HERRING

Synopsis of Panel Review

The Panel reached consensus on all terms of reference for this stock.

There was a rigorous analysis and review of all datasets considered for the assessment model with good justification given for the inclusion or exclusion of each. One of the datasets, that of the acoustic survey, received special attention and based on issues such as coverage, was not used in the assessment. The Panel had concerns that it was not possible to include uncertainty in some of the survey calibration coefficients (e.g. Bigelow to Albatross) in the assessment model. As well, the magnitude of the NMFS survey catchability increase due to the trawl door change in 1985 was very different for the spring and fall surveys, which requires further examination.

The Panel accepted the new ASAP base case assessment model as the most plausible to inform management decisions. The new model employed a new age and time-specific natural mortality (M), but the Panel, while accepting the increase in the long-term average M from 0.2 to 0.3, considered that the application of an age and time-specific pattern based on Lorenzen (1996) went beyond the data and was unnecessary. The key feature of this new assessment model is the 50% increase in natural mortality over the long-term average (0.3) since 1996, which is consistent with data on herring predation and largely resolves the retrospective pattern which has been a prominent feature of previous assessments. The Panel highlights the significant work undertaken by the HWG to document long-term trends in herring consumption by predators which allowed consideration of the increased M hypothesis. The biological reference points were derived assuming that the 50% increase in M due to herring consumption will continue over the next 3 – 5 years. Monitoring is required to determine whether or not this increase will continue over the longer term and if the reference points will continue to be appropriate.

Another source of uncertainty in the stock projections is the size of the 2008 year-class, which has been assessed as being almost twice as large as the next largest (1994). This year-class is prominent in the recent stock biomass increase and will be a significant component of projected yield over the immediate future. It will be important to monitor the size of this year-class to confirm its strength. Notwithstanding the uncertainties, the Panel concurred with the HWG that the stock is not overfished and that overfishing is not occurring.

The Panel considered that the response of the management system to uncertainties caused by high recent M and other processes is most appropriately examined using Management Strategy Evaluation. Some work on aspects of this has been undertaken; the Panel strongly encourages its further pursuit.

Evaluation of Terms of Reference

1. *Estimate catch from all sources including landings and discards. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.*

The Panel concluded that this term of reference had been successfully completed.

The spatial distribution of fishery landings by month from 2006 to 2011 was described in detail. US catch data from fixed and mobile fishing gears and catch data from the New Brunswick weir fishery were collated to produce time series of catch data from the mobile and fixed gear fleets for the Gulf of Maine/Georges Bank Atlantic herring stock complex. Given the decision on stock boundary, the HWG has done a good job in compiling catch landings and discards within the stock boundary separated into fixed and mobile gear types.

Length and age sampling was extensive except for US fixed gear in more recent years. There were major differences in length frequencies shown by gear type, but not for the same gear in different areas, so the data were not compiled by individual areas. There was clear evidence of strong and weak year classes in the composition data, giving some indication of good data quality.

Recent observer coverage has been in the order of 30% overall, 100% in closed area 1 and 80% on Georges Bank. The coverage rate 10 years ago was 5-15%. Discards are less than 1% of landings, and observer coverage is adequate to give confidence in that figure.

Some account of uncertainty in catch was made for each fleet through the CV on total landed catch, and tuning of the effective sample size on proportions at age. However, greater uncertainty of catches from the stock may be due to the decision on placement of the stock boundary. In particular, the influence of catches in the Scotian Shelf region on the stock is unknown. It would be possible to develop alternative catch series that take some account of the uncertainty in the stock boundary. Such alternatives were not developed, so the current assessment does not fully consider uncertainty in catch, discards, or the spatial distribution of the fishery.

2. *Present the survey data being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, larval surveys, age-length data, predator consumption rates, etc.). Investigate the utility of commercial LPUE as a measure of relative abundance, and characterize the uncertainty and any bias in these sources of data.*

The Panel concluded that this term of reference has been successfully completed.

Abundance indices of Atlantic herring derived from the NMFS spring, fall, winter, and summer shrimp surveys, a larval index derived from NMFS ichthyoplankton surveys, and abundance indices derived from the Massachusetts Division of Marine Fisheries and the joint Maine/New Hampshire bottom trawl surveys were presented and the utility of each index was examined. The Panel noted that indices from the winter survey were not employed in the assessment because of inconsistent spatial coverage and lack of fit. The larval index was not used because the relationship between the index and spawning stock biomass was confounded by predation of herring eggs and the relationship of the larval index to recruitment of age-1 recruits was unclear. The MA DMF and ME/NH were not used due to poor coverage of the stock. The Panel endorsed the decision to use the NMFS spring, fall, and shrimp bottom trawl surveys, and not use the winter, larval, and state-run surveys.

The use of commercial LPUE as an index of abundance was discounted because of the effect of fishing regulations on locations fished, hyperstability arising from use of sonar to track schools of fish, and the difficulty of identifying "herring trips" because of target switching within trips. The Panel agreed that LPUE would not provide a useful index of abundance.

The Panel strongly endorsed the decision to stop applying commercial age-length keys to estimate age compositions from survey length composition data.

The Panel noted that the consistency of surveys had been affected by changes to nets (e.g., use of Yankee 41 for spring surveys between 1973 and 1981 rather than standard Yankee 36 net), trawl gear (e.g., change of trawl door in 1985), and survey vessels (e.g., use of FRV Delaware II rather than FRV Albatross IV for spring surveys in 1973, 1979-1982, 1989-1991, 1994, and 2003 and fall surveys in 1977-1978, 1980-1981, 1989-1991, and 1993, and change from FRV Albatross IV to FSV Henry B. Bigelow in 2009), and that uncertainty is present in the estimates of the calibration factors used to standardize survey data to FRV Albatross IV equivalent data. The value of the spring and fall surveys is reduced by the substantial uncertainties in their calibration. The Panel observed that the length-specific Bigelow-Albatross calibration for the fall survey seems unnecessarily and implausibly complex, but recognized that this is inconsequential and would be likely to have only a negligible effect on the results of the assessment. The Panel was concerned that, because of the paucity of data for Atlantic herring in paired-tow experiments, a calibration factor could not be determined to convert survey data collected using the Yankee 41 net to Yankee 36 equivalent values.

The Panel concluded that, overall, the survey indices selected for use in the model were adequate for the current assessment.

3. *Evaluate the utility of the NEFSC fall acoustic survey to the stock assessment of herring. Consider degree of spatial and temporal overlap between the survey and the stock. Compare acoustic survey results with measures derived from bottom trawl surveys.*

The Panel concluded that this term of reference had been successfully completed.

The Panel concurred with the decision not to use the NEFSC fall acoustic survey index. The marked decline in the values of the acoustic index following the first three years of the survey was inconsistent with the trends exhibited by the other indices, and the model could not explain the drop. Various hypotheses were proposed as to why the decline in the values of the index might be an artifact rather than a reflection of the true trend in abundance, e.g., difference in fish behavior and vertical distribution during the time of the acoustic survey. Back-calculated estimates of spawning time derived from data collected in the larval study provided little support for the hypothesis that the decline was due to differences in the relationship of spawning time with the time at which the survey was undertaken.

The temporal and spatial extent of survey was considered in relation to the distribution of Atlantic herring. The survey was seen to cover a limited spatial area that was not representative of the entire stock.

4. *Evaluate the validity of the current stock definition, and determine whether it should be changed. Take into account what is known about migration among stock areas.*

The Panel concluded that this term of reference had been successfully completed.

There are at least three major sub-stocks of herring within the defined boundaries of the management region encompassing the Gulf of Maine / Georges Bank stock complex with known mixing at the northern and southern boundaries. Given the requirement for an assessment of the stock, one major sub-stock confined within the Gulf of Maine/Georges Bank region and uncertainty in the level of mixing of the other two sub-stocks, the Panel agrees with the HWG decision to maintain the current stock definition for management purposes.

Although the HWG has stated that separation of catches and catch composition information by sub-stocks is not possible at present, the Panel believes that such data separation is one of the major barriers to improving the stock assessment, and agrees with the HWG that future research should be directed towards data separation by sub-stock. For example, movement rates from tagging studies may be used to create generalized sub-stock mixing rates which could then be used in a multi-substock assessment, or used directly to separate assessment input data by sub-stock.

Sub-stock mixing at the southern and particularly northern stock boundaries introduces one of the major uncertainties into the stock assessment. Scenarios should be developed that account for this uncertainty that can be carried through to the stock assessment. One such approach is to develop alternative catch scenarios as discussed under term of reference 1.

5. *Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-6), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.*

The Panel concluded that this term of reference was successfully completed.

It accepted the base case assessment model as the most plausible model to inform management decisions. The key features of the base case model are a long-term average natural mortality (M) over age of 0.3 and a 50% increase in this M since 1996. It is rare that such an hypothesis of increased M is used in stock assessments but the temporal trend in predator consumption of herring and the positive effect on retrospective patterns justifies its use.

The assessment model underwent many significant changes since the 2009 assessment (TRAC, 2009). The continued use of the ASAP forward-projecting statistical catch at age model is supported by the Panel. The key innovation in the current formulation is the change to M, which is undertaken in two components. The first is age and time-specific and is related to life history and is applied to the assessment's time series (1965-2011). The second is associated with the consumption of herring by predators (see term of reference 6) and is additional to the long-term M during 1996 – 2011.

The long-term average M was estimated (0.3) by the method of Hoenig (1983) modified by an age and time-specific pattern estimated from weight at age by the method of Lorenzen (1996). The Panel noted that the latter method was intended to describe broad trends in ecosystems between species body weight and natural mortality and is ill-suited to the current task. It noted that the HWG had conducted assessment model runs, which had indicated that use of an age and time specific pattern in M provided similar results as using a fixed M based on Hoenig (1986) alone. The Panel considers that use of an age and time-specific pattern in M goes beyond the resolution of the data and is inconsequential to the assessment.

The use of a post-1995 increased M is consistent with the herring predation data and largely resolves the retrospective pattern that has been a prominent feature of previous assessments. During the review, the change in model likelihood, Mohn's Rho and average (GM) 1996 – 2010 consumption was estimated for a range of percent increases in post-1995 M. This indicated that the 50% increase used in the base model was appropriate and plausible.

The Panel noted that a profile of likelihoods over the steepness of the Beverton and Holt (BH) stock – recruitment relationship indicated the low precision in this parameter, with a 95% confidence interval ranging from less than 0.35 to 0.85. The model indicates a recent increase in stock biomass, largely driven by the large 2008 year-class, estimated to be almost twice as large as the next largest (1994) year-class. During the review, an exploratory run was undertaken with the CV on the recruitment deviations about the BH relationship set to that produced by the base model. While the size of the 2008 year-class was reduced somewhat, it still remained by far the strongest in the assessment time series.

The Panel considers that the uncertainties in the assessment were well characterized by the MCMC analyses and runs of alternate formulations. There are, however, some issues that require future attention. It was unfortunate that it was not possible to include uncertainties in the survey calibrations (e.g. Bigelow to Albatross) in the assessment. There was also a concern on the extent of the NMFS spring and fall survey catchability (q) increase (2.64 for spring and 13.6 for fall) due to a trawl door change in 1985. It is difficult to understand why the q change would be so different between spring and fall and the mechanism that would cause this.

6. *Consider the implications of consumption of herring, at various life stages, for use in estimating herring natural mortality rate (M) and to inform the herring stock-recruitment relationship. Characterize the uncertainty of the consumption estimates. If possible integrate the results into the stock assessment.*

The Panel concluded that this term of reference had been, for the most part, successfully completed. Uncertainty of the consumption estimates for fish and highly migratory species (HMS) predation was not discussed in the assessment report.

Estimates of consumption of Atlantic herring were used to inform the decision to increase post-1995 natural mortality in the base assessment model. The Panel noted that, although it is unusual to consider annual consumption as a mechanism for adjustment of natural mortality in a stock assessment, it was justified in this case because of the wealth of available annual information from stomach contents for a wide range of herring predators (fish, mammals, other), and annual abundance estimates of those predators.

The information on annual consumption of Atlantic herring was obtained from a comprehensive examination of predator and associated consumption data. The predators comprised 13 fish species, which accounted for 97% of all occurrences of herring in gut contents, two highly migratory species, which are the primary large pelagic predators of herring in the region, marine mammals, and sea birds. The total annual herring consumption by each predator was estimated from a combination of data on consumption per predator, diet composition, and predator abundance. Examination of the resulting consumption estimates indicated that most of the signal in the annual consumption of herring was derived from stomach contents rather than predator abundance. However, although uncertainty was mentioned in the sections of the assessment report relating to predation by marine mammals and sea birds, the uncertainty of the consumption estimates for fish and HMS predation was not discussed.

Despite high uncertainty in the resulting estimates of annual herring consumption, the data provided good evidence that consumption of herring had increased since 1996. The Panel noted, however, that (1) the later peak in the time series of annual consumption estimates was driven by the very high abundances of two individual predator species, but similar high levels of abundance for those species were not present in adjacent years; (2) abundance estimates of some predator species were calculated from swept area calculations, rather than assessment models; (3) the consumption estimates used in this assessment were likely to be underestimates of total consumption; and (4) estimates of consumption were included in an exploratory run of the assessment model but did not improve model fit.

Based on the advice that was presented during the review meeting, the Panel agreed that herring consumption was likely to remain high in the immediate future given the trends in the abundance of predator populations.

7. *State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.*

The Panel concluded that this term of reference had been successfully completed.

The Panel endorsed the redefinition of the biological reference points to those derived using the base model from this assessment.

The Panel noted that the updated value that defines the overfished level of spawning stock biomass (SSB) is considerably less than the existing value of this reference point. This is due to the fact that existing biological reference points had been derived using an age-aggregated rather than age-structured model, and had been based on an analysis of catch data aggregated over gear types rather than an analysis that employed catches by mobile or fixed gear. The updated reference points were analytically derived using the parameters of a BH stock-recruitment relationship fitted internally in the base assessment model.

The sensitivity of the BRPs to uncertainty associated with age-specific M, various levels of percentage increase in M since 1995, and steepness of the BH stock-recruitment relationship were examined by the Panel. The Panel concluded that the BRPs calculated using the base model, i.e., with a base level of M = 0.3 and a post-1995 increase in M of 50% due to consumption by predators, were appropriate for the immediate future (3 to 5 years). The Panel recommended, however, that monitoring of predation be continued due to large uncertainty in the assumption that consumption of Atlantic herring by predators would remain at current levels in the longer-term.

8. *Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model, should one be developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).*
    - *When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.*
    - *Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-7).*

The Panel concluded that this term of reference was successfully completed.

It concurs with the HWG's conclusion that, while there is uncertainty in the reference points, the stock is not overfished and that overfishing is not occurring.

Current status was determined using the previous model assumptions as not overfished and overfishing was not occurring.

Status was determined using the new model and updated data under a range of M (both age and time – specific and post-1995) and BH steepness options and, except for the unlikely case where steepness was assumed to be 0.35, the current status was estimated to be not overfished and overfishing was not occurring. It was reassuring that, although resulting biological reference points varied greatly in response to the changes in steepness, conclusions regarding current stock status remained unchanged from that of the base model except in the (rather unlikely) case of a steepness of 0.35, where overfishing (but not overfished) was indicated.

9. *Using simulation/estimation methods, evaluate consequences of alternative harvest policies in light of uncertainties in model formulation, presence of retrospective patterns, and incomplete information on magnitude and variability in M*

The Panel concluded that this term of reference was not completed, but that some initial work was underway.

The task described in this term of reference is very substantial, and probably more than can be accomplished within the timeframe of a stock assessment. Further, it is a task that cannot be usefully completed by scientists alone. It requires close engagement and extensive consultation with both managers and stakeholders in order to elicit the quantifiable management objectives and performance measures that would provide the framework for the simulation study. Such consultation would establish what sort of harvest policies could be acceptable to managers and stakeholders, and some of the criteria by which they should be evaluated. Then, these policies must be structured as formal harvest control rules so as to be incorporated in a simulation study. The assessment report briefly mentions several research projects which relate to this term of reference, but these all focus on assessment problems (e.g., retrospective patterns and mis-specification of natural mortality) rather than alternative harvest policies that might be robust to them.

The Panel considers that the study described in this term of reference could be of great help in the management of Atlantic herring, particularly considering the uncertainty about how long the current high level of herring consumption may persist.

10. Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

- *Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).*
- *Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.*
- *Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.*

The Panel concluded that this term of reference was successfully completed.

The projection methods were sound, and were applied to a wide range of scenarios that successfully spanned the plausible range of uncertainty. Key sources of this stock's vulnerability to becoming overfished (e.g., contributions from other herring stocks, and uncertainty about the strength of the 2008 year class and the persistence of high natural mortality) were identified and well described.

11. For any research recommendations listed in recent peer reviewed assessment and review panel reports, review, evaluate and report on the status of those research recommendations. Identify new research recommendation

The Panel concluded that this term of reference had been successfully completed.

There were no research recommendations from previous assessments.

The HWG compiled a new list of recommendations as part of the most recent assessment. Those recommendations can be assigned to groupings: (1) improve incorporation of sub-stocks in the assessment (a,b,c); (2) improve diet composition information (d,j,k,l,m); (3) examine novel data collection procedures or indices (e,i,n); and (4) examine improved assessments methods (g,h,o). In terms of priority, the herring assessment has the most scope for improvement through additional support for the M adjustment (group 2) and to better account for sub-stock structure (group 1).

The Panel suggests that recommendation f could be modified to "Determine the depth distribution of Atlantic herring and the factors that influence that distribution", and that recommendation g should be made more specific.

In addition, the Panel recommends:

- Alternative catch scenarios be developed to account for uncertainty in the stock boundary, particularly including catches from the Scotian Shelf. This would also allow examination of whether catch underestimation (e.g. inclusion of Scotian shelf catch) can contribute to the reduction in the retrospective pattern and contribute to or explain the need for increased M.

- Look at the effect of adding a penalty to encourage the NMFS survey trawl door-change q ratios to be similar in spring and fall.

- Using simulation/estimation methods, evaluate consequences of alternative harvest policies in light of uncertainties in model formulation, presence of retrospective patterns, and incomplete information on magnitude and variability in M (see term of reference 9).

BIBLIOGRAPHY

Panel Summary Report References

Bull, B., R.I.C.C. Francis, A. Dunn, A. McKenzie, D.J. Gilbert, M.H. Smith, R. Bian and D. Fu . 2012. CASAL (C++ algorithmic stock assessment laboratory): CASAL User Manual v2.30-2012/03/21. *NIWA Technical Report 135*. 280 p.

McAllister, M.K., and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54(2): 284–300.

NEFSC. 2008. Northeast Fisheries Science Center. 2008. Assessment of 19 Northeast Groundfish Stocks through 2007: Report of the 3rd Groundfish Assessment Review Meeting (GARM III), Northeast Fisheries Science Center, Woods Hole, Massachusetts, August 4-8, 2008. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 08-15; 884 p + xvii.

Yevseyenko, S. A., and M. M. Nevinskiy. 1981. On the development of eggs and larvae of the yellowtailed dab Limanda ferruginea Storer and their passive transport in the Northwestern Atlantic. J. Ichthy., 21: 65–74.

Herring

Background Papers

Shepherd, G., Cieri, M., Power, M., and Overholtz, W. 2009. Transboundary Resources Assessment Committee Gulf of Maine/Georges Bank Atlantic Herring Stock Assessment Update. TRAC Reference Document - 2009/04. 44 p.

Transboundary Resources Assessment Committee. 2009. Gulf Of Maine- Georges Bank Herring Stock Complex. TRAC Status Report 2009/04. 6 p.

Transboundary Resource Assessment Committee. 2006. Gulf of Maine- Georges Bank Herring Stock Complex. TRAC Status Report 2006/01. 7 p.

Gavaris, S. 2003. Transboundary Resource Assessment Committee (TRAC): Report of Meeting held 10-14 February 2003. DFO Canadian Science Advisory Secretariat (CSAS).TRAC Proceed. 2003/014. 27 p.

O'Boyle, R. and Overholtz, W. 2006. Proceedings of the Transboundary Resources Assessment Committee (TRAC). Benchmark Review of Stock Assessment Models for Gulf of Maine and Georges Bank Herring 2 – 5 May 2006, Woods Hole, Massachusetts. TRAC Proceed. 2006/01. 31 p.

O’Brien, L., and T. Worcester, T. 2009. Proceedings of the Transboundary Resources Assessment Committee (TRAC): Gulf of Maine/Georges Bank Herring, Eastern Georges Bank Cod and Haddock, Georges Bank Yellowtail Flounder. Report of Meeting held 8 - 11 June 2009, St. Andrews Biological Station, St. Andrews, New Brunswick, Canada. TRAC Proceed. 2009/01. 38 p.

Jech, J. M. and Stroman, F. 2012. Aggregative patterns of pre-spawning Atlantic herring on Georges Bank from 1999-2010. *Aquat. Living Resour.*, 25: 1–14 (2012). 14 p.

Makris, N. C., Ratilal, P., Symonds, D. T., Jagannathan, S., Lee, S. and Nero, R. W. 2006. Fish population and behavior revealed by instantaneous continental shelf–scale imaging. *Science*, 311: 660-663.3 p.

Makris, N. C., Ratilal, P., Jagannathan, S., Gong, Z., Andrews, M., Bertsatos, I., Godø, O. R., Nero, R. W. and Jech, J. M. 2009. Supporting online material for: Critical population density triggers rapid formation of vast oceanic fish shoals. *Science*, 323: 1734. 9 p.

Link. J. S. and Almeida, F. P. 2000. An Overview and History of the Food Web Dynamics Program of the Northeast Fisheries Science Center, Woods Hole, Massachusetts. NOAA Technical Memorandum NMFS-NE-159. 60 pp.

Kanwit, J. K. and Libby, D. A. 2009. Seasonal movements of Atlantic herring (*Clupea harengus*): results from a four year tagging study conducted in the Gulf of Maine and Southern New England. *J. Northw. Atl. Fish. Sci.*, 40: 29–39.11 p.

Guan, W., Cao, J., Chen, Y. and Matthew Cieri, M. 2012. A simulation study to evaluate impacts of spatial structure of Atlantic herring fishery on retrospective errors in stock assessment. Submitted. 39 p.

Reid, R. N., Cargnelli, L. M., Griesbach, S. J., Packer, D. B., Johnson, D. L., Zetlin, C. A., Morse, W. W. and Berrien, P. L. 1999. Essential Fish Habitat Source Document: Atlantic Herring, *Clupea harengus*, Life History and Habitat Characteristics. NOAA Technical Memorandum NMFS-NE-126. 48 pp.

Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology. 49: 627–647.21 p.

Miller, T.J., Richardson, D.E., Hare, J.A. 2012. Maximum likelihood estimation of larval production indices from length frequency and growth information collected on ichthyoplankton surveys. Draft. 13 p.

Cao, J., Guan, W., Chen, Y., Cieri, M. 2012. Evaluating factors influencing retrospective errors in estimating stock biomass for Atlantic herring *Clupea harengus*. 29 p.

Working Papers for Review

Working Group, Stock Assessment Workshop (SAW 54) 2012. Stock Assessment of Atlantic Herring- Gulf of Maine/Georges Bank. SAW/SARC 54. June 5-9 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 280 p.

Working Group, Stock Assessment Workshop (SAW 54). 2012. Stock Assessment of Atlantic Herring- Gulf of Maine/Georges Bank. Appendices 3-6. SAW/SARC 54. June 5-9 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA. 57 p.

Working Group, Stock Assessment Workshop (SAW 54). 2012. Atlantic Herring Assessment Summary for 2012. SAW/SARC 54 June 5-9, 2012. NOAA Fisheries, Northeast Fisheries Science Center, Woods Hole, MA. 12 p.

Yellowtail

Background Papers

Wood, A. D., Cadrin, S. X., Alade, L. A., Martins, D., Moser, J., and Westwood, A. D. 2012. Mortality and movement of yellowtail flounder, *Limanda ferruginea*, tagged off New England. Draft in review. 32 p.

McBride, R., Press, Y., and Wuenschel, M. 2012. Classifying female yellowtail flounder maturity: comparing at-sea, macroscopic maturity classifications with results from a gonad histology method. Working paper for SARC 54. 11 p.

McElroy, W. D., Towle, E. K., Press, Y. K., McBride, R. S., and Wuenschel, M. J. 2012. Comparison of fecundity among stocks of female yellowtail flounder, *Limanda ferruginea*.Working paper for SARC 54. 12 p.

McElroy, W. D., Press, Y. K., and Wuenschel, M. J. 2012. Reproductive effort as a predictor of the natural mortality rate for southern New England yellowtail flounder: the Gunderson method. Working paper for SARC 54. 5 p.

Alade, L., and Cadrin, S, 2012. A review of yellowtail flounder stock structure off New England. Working paper for SARC 54. 13 p.

Alade, L., Hart, D., and Legault, C. 2012. Influence of spatial stratification on discard estimation for the Southern New England Mid-Atlantic yellowtail flounder. Working paper for Southern Demersal Data Meeting. 14 p.

Alade. L. 2012. Southern New England-Mid Atlantic yellowtail length-weight relationship. Working paper for Southern Demersal Data Meeting. 5 p.

Barkley, A. and Cadrin, S. 2012. Results of the application of reflex action mortality predictors onboard commercial fishing vessels. Working paper for the Southern Demersal Working Group Meeting. 10 p.

Gervelis, B., Burchard, K., and Hoey, J. 2011. A comparison of discard estimates of yellowtail flounder using study fleet self-reported data and NEFOP observer data. Northeast Fisheries Science Center. Northeast Cooperative Research Program. 18 p.

Alade, L. A. 2012. Southern New England Mid-Atlantic yellowtail flounder commercial landings. Working paper for Southern Demersal Data Meeting. 25 p.

Hare, J., Richardson, D., and Mountain, D. 2012. Indices of the Mid- Atlantic cold pool and relationship to Southern New England/ Mid-Atlantic yellowtail flounder recruitment. Working paper for SARC 54. 2 p.

Richardson, D. E., Walsh, H., and Hare, J. 2012. Southern New England yellowtail flounder larval data. National Marine Fisheries Service, Oceanography Branch. 12 p.

Anonymous. 2012. Revised Spatial Stratification for estimating discards for SNEMA yellowtail flounder. Working paper. 10 p.

Previous Assessments

GARM. 2002. Southern New England yellowtail flounder. 21 p.

GARM. 2005. Southern New England-Mid Atlantic yellowtail flounder. 26 p.

GARM. 2008. Southern New England/Mid Atlantic yellowtail flounder. 36 p.

Stock Assessment Review Committee (SARC 36). 2003. Yellowtail Flounder Stock Structure. 36th SAW Advisory. 3 p.

Northeast Fisheries Science Center. 2003. Report of the 36th Northeast Regional Stock Assessment Workshop (36th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref. Doc. 03-06; 453 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026. 453 p.

Stock Assessment Review Committee (36th SARC). 2003. Southern New England - Mid Atlantic Yellowtail Flounder. 36th SARC Consensus Summary. 63 p.

Working Papers for Review

Southern Demersal Working Group, Stock Assessment Workshop (SAW 54). 2012. Southern New England Mid-Atlantic (*Limanda ferruginea*) stock assessment updated through 2011. A report of the Southern Demersal Working Group for SARC 54, June 5th – 9th, 2012. Northeast Fisheries Science Center. Woods Hole, MA. SAW/SARC 54. 226 p.

Working Group, Stock Assessment Workshop (SAW 54). 2012. Southern New England Yellowtail Flounder Assessment Summary for 2012. SAW/SARC 54. 15 p.

NIWA
Taihoro Nukurangi

Report on the 2012 assessments of yellowtail flounder and herring at Woods Hole

Prepared for The Center for Independent Experts
Northern Taiga Ventures, Inc.

July 2012

Authors/Contributors:
R.I.C.C. Francis

For any information regarding this report please contact:

R.I.C.C. Francis
Principal Scientist

+64-4-386 0525
c.francis@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
301 Evans Bay Parade, Greta Point
Wellington 6021
Private Bag 14901, Kilbirnie
Wellington 6241
New Zealand

Phone +64-4-386 0300
Fax +64-4-386 0574

NIWA Client Report No:	WLG2012-26
Report date:	June 2012
NIWA Project:	NTV13301

Contents

Reviewed by

Dr Sophie Mormede

Approved for release by

Dr Julie Hall

Executive summary

A panel met 5-9 June 2012 at the Northeast Fisheries Science Center in Woods Hole, Massachusetts to review the 2012 assessments of the stocks of Southern New England yellowtail flounder (*Pleuronectes ferrugineus*) and Atlantic herring (*Clupea harengus*) that are fished off the northeast coast of the USA. The assessments, and some additional analyses, were presented and discussed; the Assessment Summary Reports were reviewed and edited; and the panel began drafting its Summary Report.

The review process was well run. The panel was well supported, the assessment teams were ready and able to respond to queries, and helpful comments were received from other participants.

I conclude that both assessments are sound, and thus provide a scientifically credible basis for developing fishery management advice. All terms of reference for the assessments were successfully completed, with only one minor exception, which did not compromise the assessments.

The following recommendations apply to both assessments unless otherwise stated:

- User and Technical Manuals, together with input and report files from base runs should be provided to reviewers of all ASAP assessments;
- The approach to data weighting in assessment models should be formalized;
- The fixing of stock-recruit steepness in assessments should be considered;
- The use of prior distributions on catchability ratios should be considered as a means of including uncertainty about survey calibration constants;
- The use of alternative catch histories should be considered as a means of quantifying uncertainty in matters such as discards and stock boundaries;
- Alternative approaches to estimating initial depletion should be considered;
- Inferring age- or time-dependent natural mortality from somatic weight only should be avoided;
- When Bayesian estimates are available, only these should be presented;
- Ambiguity about the final year in projections should be removed;
- Spring and fall survey biomass indices for yellowtail flounder should be standardized to remove the effects of diurnal variation in catch rates;
- The fact that alternative methods of calculating yellowtail reference points produced very different results should be explored.
- A prior distribution should be used to constrain the factors by which herring survey catchability increased in 1985 (when the trawl doors were changed) to be more similar for the spring and fall surveys; and
- Explanations for the dramatic increase in herring consumption in the mid-1990s should be sought.

1 Background

This report reviews, at the request of the Center for Independent Experts (see Appendix B), the 2012 assessments of the stocks of Southern New England yellowtail flounder (*Pleuronectes ferrugineus*) and Atlantic herring (*Clupea harengus*) that are fished off the northeast coast of the USA. The author was provided with relevant documents (Appendix A), and participated both in the meeting that considered the assessments, and in the writing of the Summary Report from that meeting.

2 Review activities

The 54th Stock Assessment Review Committee (SARC 54) met 5-9 June 2012 at the Northeast Fisheries Science Center of NOAA/NMFS in Woods Hole, Massachusetts. Those attending the meeting included the four Panel members, the Chairman and Coordinator of SARC, the Chief of the Population Dynamics Branch at Woods Hole, members of the stock assessment working groups (WGs), and other interested parties from both the fishing industry and the research community (Appendix C). The assessments, and related material, were presented to the Panel, and some additional analyses requested by the Panel were carried out and discussed. The Assessment Summary Reports were reviewed and edited, and the Panel began drafting their Summary Report.

3 Findings

The review process was very well run. The Panel was well supported and ably chaired. I was impressed by the willingness, and ability, of the assessment teams to respond to panel requests, and was particularly grateful for helpful and constructive comments given by other meeting participants.

I first present findings that are common to both assessments, and then those relating to the two individual assessments. The latter findings are grouped by the Stock Assessment Terms of References (TORs), as given in Annex 2 of Appendix B.

3.1 Findings common to both assessments

The findings in this section relate primarily to the TORs concerning the stock assessment model (i.e., TOR 4 for yellowtail flounder, TOR 5 for herring) though they do also affect other TORs (e.g., on stock status and projections).

3.1.1 ASAP

I find that ASAP (Legault & Restrepo 1999), the modeling package used in the assessments of both stocks, to be an excellent tool for present purposes. It is modern (statistically-based), well-documented, and relatively simple in structure, and thus ideal for use in an environment in which many assessments have traditionally been carried out with less modern tools.

I was surprised not to find the ASAP User and Technical Manuals amongst the background papers provided to me. Also, it would have been useful if the ASAP input and report files for the base runs had been provided as appendices to the assessment reports (as is common, for example, for west coast assessments using Stock Synthesis). These files provide easy access to technical details of the assessments that can be important, but are sometimes overlooked, or difficult to find, in the main text of the assessment reports.

3.1.2 Data weighting

I found the approach to data weighting in both assessments rather ad hoc, and not well described. For example, both assessment reports mentioned the goodness-of-fit to mean age as a criterion for weighting age composition data sets, apparently in response to a recommendation of Francis (2011), but this criterion seems to have been applied qualitatively, rather than quantitatively. I think it was a mistake to use the same effective sample size for all years of each age-composition data set, because this practice ignores information about years in which sampling was particularly weak or strong (e.g., note the reference to 'extremely poor sampling in 1999' on p. 41 of the yellowtail report). I was concerned that the iterative reweighting used for the herring assessment resulted in some very high CVs (coefficients of variation) for survey indices (see Table A5-1 in the herring assessment report), which might have resulted in these data being swamped by the age composition data.

I don't believe it is possible to remove all subjective elements from data weighting in stock assessments, but I think both the present assessments would have been improved by a less ad hoc approach. As a starting point to developing more formal methods I would suggest consideration of the approach proposed by Francis (2011), which was based on the following guiding principles:

- do not let other data stop the model from fitting abundance data well;
- when weighting age or length composition data, allow for correlations; and
- do not down-weight abundance data because they may be unrepresentative.

3.1.3 Uncertainty about stock-recruit steepness

As is very common in stock assessments, the available data were not informative about the steepness of the stock-recruit relationship for either stock, so there was great uncertainty about the value of this parameter. I was surprised at the very different responses to this uncertainty in these two assessments. For yellowtail, the response was to assume no stock-recruit relationship and base biological reference points (BRPs) on a spawning potential ratio of 40%, rather than MSY; for herring, the decision was made to estimate steepness within the assessment model, and use this estimated steepness to calculate MSY-based BRPs.

Of these two responses, the former seems to me more defensible. Recent research has cast considerable doubt on our ability to estimate steepness within an assessment model (Lee et al. 2012). Further, given the very wide confidence limits implied by the steepness profile calculated for the herring assessment, it is likely that BRPs for this stock will be unstable (because estimates of steepness – and thus BRPs – may change substantially with additional data, and any changes in data weighting, in future assessments).

There is a third possible response to uncertainty in steepness, which seems better to me. That is to fix steepness, using either a default value (in New Zealand and Australia a value of 0.75 is common, for reasons given in Francis 1993), or an average from published values for the same or similar species. The effect of uncertainty about this parameter can then be evaluated by sensitivity runs with lower and higher values of steepness. Fixing steepness should make BRPs more stable.

3.1.4 Uncertainty about survey calibrations

The biomass indices from the spring and fall surveys included corrections, or calibrations, for changes in catch rates caused by various changes in vessel and gear. I was concerned that the uncertainty associated with these calibrations was not carried through into the assessment. That is understandable, because ASAP does not yet provide any mechanism to include that uncertainty. One approach that has been used in New Zealand to address this problem is to allow the user to provide prior distributions for ratios of catchabilities.

The concept is perhaps best described with an example. Consider the calibration constant of 1.22, which was used to scale up yellowtail survey indices to compensate for the change in trawl doors in 1985, and suppose that the standard error of this constant was 0.1. The idea is to split the survey time series at 1985, so that ASAP estimates a separate catchability constant, *q*, for pre- and post-1985, and to apply a normal (say) prior with mean 1.22 and standard deviation 0.1 to the ratio of these *q*s to discourage *q* estimates whose ratio differs substantially from 1.22. This involves adding a term $0.5[((q_{post}/q_{pre})-1.22)/0.1]^2$ to the objective function. An approach like this has been used in some New Zealand assessments and is implemented (in a slightly more complicated form) in CASAL (see section 6.7.5 in Bull et al. 2012).

3.1.5 Alternative catch histories

I saw scope in both assessments to use alternative catch histories as a way of evaluating the effect of uncertainties that are otherwise difficult to quantify. For yellowtail flounder, the substantial uncertainties concerning discards could be explored by constructing two alternative catch histories representing the plausible upper and lower limits in the level of discards. For herring, uncertainty about stock boundaries could be evaluated by constructing alternative catch histories which, for example, make different assumptions about catches from the Scotian Shelf.

3.1.6 Initial depletion

This finding relates to an issue that occurred to me only after the review meeting, but was not discussed in the assessment reports or during the meeting.

Because both the assessed stocks have been exploited for much longer than the periods covered by the assessment models, there is reason to expect that the stocks may have been depleted in the first year of the assessment. That is, we might expect that the initial depletion, $SSB_{initial}/SSB_0$, would be less than 1 (where $SSB_{initial}$ and SSB_0 are the estimated spawning biomasses for the first year of the assessment and the unexploited stock, respectively). In fact, this was not the case: by my calculation, the initial depletion was 1.39 for yellowtail and 0.99 for herring. Although I can't rule out the possibility that these estimates are correct, I am concerned that they may be wrong because the model structure and data did not allow good estimates of initial depletion.

There are two modifications to these assessments that might shed light on this matter. The first is to extend the assessment period closer to the beginning of these fisheries by constructing much longer catch histories. I don't know how difficult this might be, but I suspect that the initial years in the current assessments were determined by the availability of age-composition data (as is required for VPA-type assessments, but not with statistical models) rather than historic catches. With this approach we must ignore all variations in recruitment for years before the first age-composition data, and assume that $SSB_{initial} = SSB_0$ in the new (earlier) initial model year. The hope is that the additional information (the catches for earlier years) might allow the model to obtain a better estimate of depletion for the year that is currently the initial year in these assessments.

The second modification is to retain the current assessment period but make the simplifying assumption that in the years preceding the initial year, the population had reached equilibrium under a constant fishing mortality, $F_{pre\text{-}initial}$, which must be estimated (and which need not be the same as F in the initial year). With this approach, the numbers at age in the initial year (and thus $SSB_{initial}$) would be determined by the estimated values of $F_{pre\text{-}initial}$ and SSB_0. This second approach is a refinement of a suggestion I made during the review meeting, that a catch-curve analysis of the estimated initial numbers at age might be a useful diagnostic for stocks where there were substantial catches in years before the first assessment year.

3.1.7 Natural mortality

I think that both assessments made inappropriate and unnecessary use of a regression equation from Lorenzen (1996) to infer trends in *M* (natural mortality): age-dependent *M* for yellowtail; and age- and time-dependent *M* for herring.

The use of Lorenzen's equation (which predicts *M* from somatic weight) was *inappropriate* because this equation is highly imprecise when used to predict *M* for an individual species (let alone for variation in *M* within a species). This imprecision is evident in the wide scatter about the regression line in Lorenzen's Figure 1 (from this scatter I infer that predicted and observed *M*s would differ by a factor of more than 2 for about one-third of his data points). I note that for both herring and yellowtail the estimated *M*s from Lorenzen's equation were scaled down substantially (by a factor of about 3 for yellowtail) to be more consistent with other estimates (e.g., from Hoenig's method). If Lorenzen's equation is so unreliable for mean *M* for these species, why should we consider it reliable for predicting how *M* varies within these species?

The power of Lorenzen's equation (and of an analogous equation I have seen that relates longevity to somatic weight in terrestrial animals) is in inferring differences in *M* between fish or animals of very different weights (note that Lorenzen's data cover about six orders of magnitude in somatic weight, and three orders of magnitude in *M*). From the terrestrial equation I would be very happy to infer that dogs live longer than mice. But I would not want to use this equation to make any inference about the relative longevity of poodles and Labradors, and I certainly would not want to calculate how much my dog's life expectancy had increased because it had put on weight.

The use of Lorenzen's equation in these assessments was *unnecessary* because it made no contribution to goodness of fit in the assessment models. For both species, alternative models with constant *M* (apart from the step increase in 1996 for herring) fitted the age composition data just as well as the base models. This being so, Occam's razor suggests that we stick to the simpler (constant *M*) assumption. Note that if age-dependent *M* does produce a clearly better fit to the age composition data, then it is more sensible to estimate the age dependence within the assessment model.

3.3 Findings for herring

3.3.1 TOR 1: landings and discards

Estimate catch from all sources including landings and discards. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.

I conclude that this TOR was successfully met.

I support the decision to use two fleets (fixed and mobile) and thus reduce year-to-year variation in selectivity. Clear evidence of strong and weak year classes in the age composition data is a clear signal of good quality data. There seemed to be no major sources of uncertainty in these data (except for matters relating to stock definition).

3.3.2 TOR 2: survey data

Present the survey data being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, larval surveys, age-length data, predator consumption rates, etc.). Investigate the utility of commercial LPUE as a measure of relative abundance, and characterize the uncertainty and any bias in these sources of data.

I conclude that this TOR was successfully met.

I commend the decision to stop applying commercial age-length keys to survey lengths, and agree with the decisions to use the shrimp survey, but not the winter, larval, and state-run surveys, in the assessment. The length-specific Bigelow-Albatross calibration function used for the fall survey (see Table A2-1 in the assessment report) seems unnecessarily and implausibly complex, though I doubt that this is consequential (i.e., changing to a simpler, and thus more plausible, function is not likely to change the assessment much). It is of concern that it was not possible to provide a calibration for the changes between the Yankee 36 and Yankee 41 nets, since the corresponding calibration constant used for yellowtail (1.75) was substantially different from 1.

3.3.3 TOR 3: the acoustic survey

Evaluate the utility of the NEFSC fall acoustic survey to the stock assessment of herring. Consider degree of spatial and temporal overlap between the survey and the stock. Compare acoustic survey results with measures derived from bottom trawl surveys.

I conclude that this TOR was successfully met.

I agree with the WG's decision not to use the acoustic survey in the base assessment. In my view the sharp and substantial drop in the survey index after the first three surveys (see Figure A3-3 in the assessment report) is the main reason not to use it: this drop is inconsistent with other data, particularly the catch history. I note also that herring echoes extend right to the boundary of the survey area in several years (see Figure A3-1 in the assessment report), so it is quite possible that the proportion of the stock covered by this survey varied substantially from year to year.

3.3.4 TOR 4: stock definition

Evaluate the validity of the current stock definition, and determine whether it should be changed. Take into account what is known about migration among stock areas.

I conclude that this TOR was successfully met.

The stock structure of herring in this area is clearly complex. Though there appear to be several spawning stocks, with complex movement patterns that are only partially known, it is often not possible to confidently allocate commercial or survey catches (or echoes) to specific spawning stocks. In these circumstances stock definitions for assessment purposes must of necessity be pragmatic compromises, rather than scientifically precise. I saw no evidence supporting a change in the existing stock definition.

3.3.5 TOR 5: the assessment

Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-6), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

I conclude that this TOR was successfully met.

I support the WG's choice of base model, and specifically the hypothesis that *M* (natural mortality) increased by 50% in the mid-1990s. It is rare that such a hypothesis can be justified in an assessment model, but the strong temporal trend in the consumption data, (TOR 6) and the fact that the increase in *M* removed a strong retrospective pattern, justify it in this case. Analyses carried out during the review provided additional support for the use of a value of 50% for the increase in *M.* The uncertainty in the assessment was well characterized by the use of MCMCs and alternative model runs.

It is a concern that the factor by which survey catchability was estimated to increase when the trawl doors were changed in 1985 was so different in the fall and spring surveys (the estimated factors were 13.6 and 2.64, respectively). Such a large difference seems implausible.

The assessment used a high recruitment CV in order that estimates of recruitment deviates be 'unconstrained' (p. 158 in the Assessment Report). I think this was a mistake (it's certainly contrary to the common practice of making this CV similar to that of the estimated recruitment deviates). This CV helps define the Bayesian prior distribution for the recruitment deviates, and it is precisely the function of such priors to constrain estimates in cases where the data are relatively uninformative. However, an alternative run (during the review meeting) with a lower CV showed that the assessment results – including the size of the very strong 2008 cohort, which is influential in projections (see TOR 10) – were not very sensitive to this CV.

The WG devised two responses to difficulties caused by occasional high proportions of age 1 fish: the removal of all age 1 fish from survey indices and age compositions; and the reduction of effective sample sizes for the fixed fleet from 29 (for 1965-1994) to 5 (post-1994). I strongly support the intent of these measures (to make the assessment robust to occasional outliers) but am concerned that the baby might have been thrown out with the bath water (the first measure removes *all* age 1 fish from survey data because of problems with *some* of them; the second down-weights *all* post-1994 fixed-fleet composition data because of problems with just one age). I wonder whether the sought-after robustness could have been achieved by using (a) a lower recruitment CV to constrain estimated recruitment deviates (see preceding paragraph), and (b) a more robust likelihood (some examples of robust likelihoods for proportions data are given in Section 6.7.1 of Bull et al. 2012).

3.3.6 TOR 6: herring consumption

Consider the implications of consumption of herring, at various life stages, for use in estimating herring natural mortality rate (M) and to inform the herring stock-recruitment relationship. Characterize the uncertainty of the consumption estimates. If possible integrate the results into the stock assessment.

I conclude that this TOR was successfully met.

The estimates of herring consumption, although very uncertain, were of great value in this assessment because their dramatic increase in the mid-1990s (see Figure A6-5 in the assessment report) was crucial in the formulation of the base model (see TOR 5).

I think further research aimed at trying to find the primary cause(s) of this dramatic increase could be useful. An analysis carried out during the review meeting showed that the increase was not caused by an increase in predator abundance. In this analysis, the annual estimates of the abundance of each predator were replaced by their time averages, and this was shown to have comparatively little effect on the time trend in annual herring consumption. I wonder whether similar analyses (i.e., the replacement of annual values with their time averages), applied to other factors that are used in calculating annual herring consumption, might be revealing. Presumably the dramatic increase described above was caused by prey switching (towards herring) by some or all predators. If we knew which prey, or group of prey, the predators switched away from, we might be in a better position to detect any future change back to pre-1990s levels of herring consumption.

3.3.7 TOR 7: stock status definitions

State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for B_{MSY}, $B_{THRESHOLD}$, F_{MSY} and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

I conclude that this TOR was successfully met.

I endorse the WG's decision to calculate BRPs using the increased value of *M* assumed in the assessment for recent years (see TOR 5). The uncertainty in BRPs was well characterized using a range of alternative assumptions (e.g., concerning *M* and steepness)

3.3.8 TOR 8: current stock status

Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model, should one be developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).

a. **When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.**
b. **Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-7).**

I conclude that this TOR was successfully met.

I endorse the WG's conclusion that the herring stock is not overfished and that overfishing is not occurring. This conclusion was shown to be robust to a wide range of alternative assumptions.

3.3.9 TOR 9: alternative harvest policies

Using simulation/estimation methods, evaluate consequences of alternative harvest policies in light of uncertainties in model formulation, presence of retrospective patterns, and incomplete information on magnitude and variability in M.

I conclude that that this Term of Reference was not completed, but that some useful initial work was underway.

The task described in this Term of Reference seemed to me to be very substantial – particularly as no guidance was given as to what types of harvest policies to consider – and out of place in a project that is primarily aimed at a stock assessment. We have had some success with this sort of study in New Zealand (particularly for rock lobster fisheries, e.g., Breen & Kim 2006), but that has required close engagement and extensive consultation with both fishery managers and stakeholders to ensure that the results are relevant to management. I note that the relevant research projects described as underway in the assessment report all focus on assessment problems (e.g., retrospective patterns and miss-specification of natural mortality) rather than the effect of these problems on harvest policies.

3.3.10 TOR 10: projections

Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

a. **Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).**
b. **Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.**
c. **Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.**

I conclude that this TOR was successfully met.

I found that the WG's projection methods were sound, and was pleased to see them applied to a wide range of scenarios so as to cover the major uncertainties in the assessment. The WG identified and described the key sources of this stock's vulnerability to becoming overfished (e.g., contributions from other herring stocks, and uncertainty about the strength of the 2008 year class and the persistence of high natural mortality).

3.3.11 TOR 11: research recommendations

For any research recommendations listed in recent peer reviewed assessment and review panel reports, review, evaluate and report on the status of those research recommendations. Identify new research recommendations.

I conclude that this TOR was successfully met.

There were no recommendations from previous assessments, but the WG presented a long list of new recommendations, all of which have some merit. This list is currently written rather tersely, which is perfectly adequate as an aide-memoire to those doing the current assessment, but which may not be very clear when considered by those doing the next assessment to be reviewed.

I would endorse recommendation h (Evaluate use of length-based models) if this is intended to refer to models that can use both length- and age-composition data. However, the term 'length-based' is more usually used to refer to models devised for stocks where no age data are available (so the models keep track only of numbers at length, rather than numbers at age – see, e.g., Kristensen et al. 2006). Such models cannot use age-composition data, and so are inappropriate for this stock.

I have two research recommendations. The first is to evaluate the use of a prior distribution to constrain the factors by which survey catchability increased in 1985 (when the trawl doors were changed) to be more similar for the spring and fall surveys (see TOR 5). This is precisely the sort of situation for which Bayesian priors are intended: there are no data directly relating to how different the factors should be for the two surveys, but expert opinion could be used to define a plausible range, and thus a prior distribution, for their ratio My second recommendation is to seek explanations for the dramatic increase in herring consumption in the mid-1990s (as discussed above under TOR 6).

4 Conclusions and recommendations

I conclude that the assessments of yellowtail flounder and herring are generally sound, and thus provide a scientifically credible basis for developing fishery management advice. All terms of reference for the assessments were successfully completed, with only one minor exception. The exception was TOR 9 for herring, and the failure to complete this TOR was understandable (given the very substantial task involved) and did not compromise the assessment of this stock.

4.1 Recommendations common to both assessments

I have nine recommendations for future assessments that are not specific to either stock (they derive directly from findings in Sections 3.1.1–3.1.9, where more detail is provided):

- User and Technical Manuals, together with input and report files from base runs should be provided to reviewers of all ASAP assessments;
- The approach to data weighting in assessment models should be formalized;
- The fixing of stock-recruit steepness in assessments should be considered;
- The use of prior distributions on catchability ratios should be considered as a means of including uncertainty about survey calibration constants;
- The use of alternative catch histories should be considered as a means of quantifying uncertainty in matters such as discards and stock boundaries;
- Alternative approaches to estimating initial depletion should be considered;
- Inferring age- or time-dependent natural mortality from somatic weight only should be avoided;
- When Bayesian estimates are available, only these should be presented; and
- Ambiguity about the final year in projections should be removed.

4.2 Recommendations for individual assessments

For future yellowtail flounder assessments I recommend that

- Spring and fall survey biomass indices be standardized to remove the effects of diurnal variation in catch rates; and
- The fact that alternative methods of calculating reference points produced very different results should be explored.

For future herring assessments I recommend that

- The use of a prior distribution to constrain the factors by which survey catchability increased in 1985 (when the trawl doors were changed) to be more similar for the spring and fall surveys; and
- Explanations for the dramatic increase in herring consumption in the mid-1990s should be sought.

5 References

Breen, P.A.; Kim, S.W. (2006). Development of an operational management procedure (decision rule) for CRA 4. *New Zealand Fisheries Assessment Report 2006/53.* 46 p.

Bull, B.; Francis, R.I.C.C.; Dunn, A.; McKenzie, A.; Gilbert, D.J.; Smith, M.H.; Bian, R. (2012). CASAL (C++ algorithmic stock assessment laboratory): CASAL User Manual v2.30-2012/03/21. NIWA Technical Report 135. 280 p.

Francis, R.I.C.C. (1993). Monte Carlo evaluation of risks for biological reference points used in New Zealand fishery assessments. *Canadian Special Publication of Fisheries and Aquatic Sciences 120*: 221-230.

Francis, R.I.C.C. (2011). Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences 68*: 1124-1138.

Kristensen, K.; Lewy, P.; Beyer, J.E. (2006). How to validate a length-based model of single-species fish stock dynamics. *Canadian Journal of Fisheries and Aquatic Sciences 63*(11): 2531-2542.

Lee, H-H; Maunder, M.N.; Piner, K.R.; Methot, R.D. (2012). Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models? *Fisheries Research 125–126*: 254– 261.

Legault, C.M.; Restrepo, V.R. (1999). A flexible forward age-structured assessment program. *ICCAT Coll. Vol. Sci. Pap. 49(2):* 246-253.

Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. *Journal of Fish Biology 49*: 627-647.

Myers, R.A. (1998). When do environment-recruitment correlations work? *Reviews in Fish Biology and Fisheries 8*: 283-305.

Appendix A Bibliography of materials provided for review

Yellowtail Flounder

Background Papers

1 Wood, A. D., Cadrin, S. X., Alade, L. A., Martins, D., Moser, J., & Westwood, A. D. 2012. Mortality and movement of yellowtail flounder, *Limanda ferruginea*, tagged off New England.

2 McBride, R., Press, Y., & Wuenschel, M. 2012. Classifying female yellowtail flounder maturity: comparing at-sea, macroscopic maturity classifications with results from a gonad histology method.

3 McElroy, W. D., Towle, E. K., Press, Y. K., McBride, R. S., & Wuenschel, M. J.. 2012. Comparison of fecundity among stocks of female yellowtail flounder, *Limanda ferruginea*.

4 McElroy, W. D., Press, Y. K., & Wuenschel, M. J. 2012. Reproductive effort as a predictor of the natural mortality rate for southern New England yellowtail flounder: the Gunderson method.

5 Alade, L., & Cadrin, S, 2012. A review of yellowtail flounder stock structure off New England.

6 Alade, L., Hart, D., & Legault, C. 2012. Influence of spatial stratification on discard estimation for the Southern New England Mid-Atlantic yellowtail flounder.

7 Alade. L. 2012. Southern New England-Mid Atlantic yellowtail length-weight relationship.

8 Barkley, A., & Cadrin, S. 2012. Results of the application of reflex action mortality predictors onboard commercial fishing vessels.

9 Gervelis, B., Burchard, K., & Hoey, J. 2011. A comparison of discard estimates of yellowtail flounder using study fleet self-reported data and NEFOP observer data.

10 Alade, L. A. 2012. Southern New England Mid-Atlantic yellowtail flounder commercial landings.

11 Hare, J., Richardson, D., & Mountain, D. 2012. Indices of the Mid-Atlantic cold pool and relationship to Southern New England/ Mid-Atlantic yellowtail flounder recruitment.

12 Richardson, D. E., Walsh, H., & Hare, J. 2012. Southern New England yellowtail flounder larval data.

13 Anonymous. 2012. Revised Spatial Stratification for estimating discards for SNEMA yellowtail flounder.

Previous Assessments

14 GARM. 2002. Southern New England yellowtail flounder.

15 GARM. 2005. Southern New England-Mid Atlantic yellowtail flounder.

16 GARM. 2008. Southern New England/Mid Atlantic yellowtail flounder.

17 Anonymous. 2003. Yellowtail Flounder Stock Structure. 36th SAW Advisory.

18 Northeast Fisheries Science Center. 2003. Report of the 36th N0ortheast Regional Stock Assessment Workshop (36th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref. Doc. 03-06; 453 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026.

19 Anonymous. 2003. Southern New England - Mid Atlantic Yellowtail Flounder. 36th SARC Consensus Summary.

Working Papers

1 Southern Demersal Working Group, Stock Assessment Workshop (SAW 54). 2012. Southern New England Mid-Atlantic (*Limanda ferruginea*) stock assessment updated through 2011. A report of the Southern Demersal Working Group for SARC 54, June 5th – 9th, 2012. Northeast Fisheries Science Center. Woods Hole, MA. SAW/SARC 54.

2 Anonymous. 2012. Southern New England Yellowtail Flounder Assessment Summary for 2012. SAW/SARC 54.

Atlantic Herring

Background Papers

1 Shepherd, G., Cieri, M., Power, M., & Overholtz, W. 2009. Transboundary Resources Assessment Committee Gulf of Maine/Georges Bank Atlantic Herring Stock Assessment Update. TRAC Reference Document 2009/04.

2 Transboundary Resources Assessment Committee. 2009. Gulf Of Maine-Georges Bank Herring Stock Complex. Status Report 2009/04.

3 Transboundary Resource Assessment Committee. 2006. Gulf of Maine-Georges Bank Herring Stock Complex. Status Report 2006/01.

4 Gavaris, S. 2003. Transboundary Resource Assessment Committee (TRAC): Report of Meeting held 10-14 February 2003. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2003/014.

5 O'Boyle, R., & Overholtz, W. 2006. Proceedings of the Transboundary Resources Assessment Committee (TRAC). Benchmark Review of Stock Assessment Models for Gulf of Maine and Georges Bank Herring 2 – 5 May 2006 Woods Hole, Massachusetts.

6 O'Brien, L., & T. Worcester, T. 2008. Proceedings of the Transboundary Resources Assessment Committee (TRAC): Gulf of Maine/Georges Bank Herring, Eastern Georges Bank Cod and Haddock, Georges Bank Yellowtail Flounder. Report of Meeting held 8 - 11 June 2009 St. Andrews Biological Station, St. Andrews, New Brunswick, Canada.

7 Jech, J. M., & Stroman, F. 2012. Aggregative patterns of pre-spawning Atlantic herring on Georges Bank from 1999-2010. *Aquat. Living Resour.*, **25**: 1–14.

8 Makris, N. C., Ratilal, P., Symonds, D. T., Jagannathan, S., Lee, S., & Nero, R. W. Fish population and behavior revealed by instantaneous continental shelf–scale imaging. *Science*, **311**: 660-663.

9 Makris, N. C., Ratilal, P., Jagannathan, S., Gong, Z., Andrews, M., Bertsatos, I., Godø, O. R., Nero, R. W., & Jech, J. M. 2009. Supporting online material for "Critical population density triggers rapid formation of vast oceanic fish shoals". *Science*, **323**: 1734.

10 Link. J. S., & Almeida, F. P. 2000. An Overview and History of the Food Web Dynamics Program of the Northeast Fisheries Science Center, Woods Hole, Massachusetts. NOAA Technical Memorandum NMFS-NE-159. 60 pp.

11 Kanwit, J. K., & Libby, D. A. 2009. Seasonal movements of Atlantic herring (*Clupea harengus*): results from a four year tagging study conducted in the Gulf of Maine and Southern New England. *J. Northw. Atl. Fish. Sci.*, **40**: 29–39.

12 Guan, W., Cao, J., Chen, Y., & Matthew Cieri, M. 2012. A simulation study to evaluate impacts of spatial structure of Atlantic herring fishery on retrospective errors in stock assessment.

13 Reid, R. N., Cargnelli, L. M., Griesbach, S. J., Packer, D. B., Johnson, D. L., Zetlin, C. A., Morse, W. W. & Berrien, P. L. 1999. Essential Fish Habitat Source Document: Atlantic Herring, *Clupea harengus*, Life History and Habitat Characteristics. NOAA Technical Memorandum NMFS-NE-126. 48 pp.

14 Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology. 49: 627–647.

15 Miller, T. J., Richardson, D. E., & Hare, J. A. In prep. Maximum likelihood estimation of larval production indices from length frequency and growth information collected on ichthyoplankton surveys.

16 Cao, J., Guan, W., Chen, Y., & Cieri, M. In prep. Evaluating factors influencing retrospective errors in estimating stock biomass for Atlantic herring *Clupea harengus*.

Working Papers

1 Anonymous. 2012. Stock Assessment of Atlantic Herring- Gulf of Maine/Georges Bank. SAW/SARC 54. June 5-9 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

1 Appendices 3-6

Anonymous. 2012. Stock Assessment of Atlantic Herring- Gulf of Maine/Georges Bank. Appendices 3-6. SAW/SARC 54. June 5-9 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

2 Anonymous. 2012. Atlantic Herring Assessment Summary for 2012.

Report on the SARC Review of SAW 54 Stock Assessments for Southern New England yellowtail flounder and Atlantic herring, June 2012

N. G. Hall

Unit 2
2 Wexford Street
Subiaco
Western Australia 6008
Australia
Phone: +61 8 9401 6891
Email: nghall@iinet.net.au

June 2012

Contents

1. Executive Summary

The SARC for the 54th Stock Assessment Workshop (SAW 54) met at Woods Hole in June 2012 to examine the stock assessments developed for Southern New England yellowtail flounder (*Pleuronectes ferrugineus*) and Atlantic herring (*Clupea harengus*). Overall, the SARC Review Panel found the assessments to be of a high quality. Decisions relating to stock definition, data to be employed, model structure, base models, and Biological Reference Points (BRPs), and findings on stock status, short term projections, and vulnerabilities by each Working Group (WG) appeared sound and were accepted and endorsed by the Panel. The Panel complimented the yellowtail flounder WG for the bridging analyses that they had conducted to determine the implications of the change from the existing Virtual Population Analysis (VPA) assessment model to a statistical catch-at-age ASAP model. Specific areas of concern in the assessments, which were explored during the review, are discussed below.

As recognised by the WG, the stock definition for the Gulf of Maine/Georges Bank (GoM/GB) Atlantic herring defines a "stock complex", which comprises a number of component sub-stocks. The definition is appropriate for the assessment given the current inability to distinguish fish from the different sub-stocks in commercial and survey catches. There is a risk, however, that stock assessments based on this stock definition may fail to detect whether less-productive stock components are overfished or being subjected to overfishing. Continued research on methods of distinguishing individual fish from component sub-stocks is urged.

For both species, there is a need to explore uncertainty of the assessments to the possibility of interchange of fish and/or larvae among the defined stock of interest and adjoining stocks. The sensitivity of the assessments to such interchange could be explored through the development of alternative catch histories and/or incorporation of alternative stock recruitment relationships, *e.g.*, use of a stock-recruitment relationship that assumes that a portion of the total recruitment arises from the spawning stock biomass (SSB) of an adjoining stock.

The decisions by the WGs to assume that natural mortality declines with weight at age in a pattern similar to that described by Lorenzen (1996), and thus varies with both age and year, were not based on available data, and, indeed, at least in the case of the GoM/GB Atlantic herring, the assumption provides no significant improvement over the simpler assumption of a constant natural mortality over age and year. A similar evaluation of the extent to which the more complex mortality assumption improves the fit for Southern New England/Mid-Atlantic (SNEMA) yellowtail flounder should be undertaken. If there is no significant improvement in the fit of the model, the more parsimonious assumption of a constant natural mortality should be adopted. Note, however, that use of the more complex assumption in the 2012 assessments is unlikely to have affected the stock status determinations that resulted from the assessments.

A key issue in the Atlantic herring assessment was whether the increase in natural mortality since 1995 that had been used in the base model was justified. The wealth of evidence on stomach contents, dietary composition and predator abundance that had been used to estimate the annual consumption of Atlantic herring by fish predators, and the resulting trends in that consumption, provided strong support for the decision that natural mortality had increased. This, combined with the improvement in retrospective patterns that resulted from the increased mortality, justified the assumption in the base model of increased mortality since 1995. An examination during the review of the change in likelihood, Mohn's Rho, and the average of the estimated consumption between 1996 and 2010 resulting from a range of percentage increases in post-1995 mortality indicated that, although uncertain, the 50% increase that had been used in the base model was appropriate. The Panel concluded that, in the immediate future, continued high levels of natural mortality were likely to occur.

A further issue in the Atlantic herring assessment was that, prior to the review meeting, aspects of uncertainty, such as the robustness of the determination of stock status to uncertainty in the magnitude of the post-1995 increase in natural mortality, had not been fully characterized. This was addressed by the various sensitivity runs of the model that were undertaken during the review meeting. The findings that the stock was not overfished and that overfishing was not occurring were robust to the uncertainties that were explored.

The key issue in the SNEMA yellowtail flounder assessment was whether the marked decline in the estimates of recruitment since 1990, which were obtained when fitting the base model, was simply the result of reduced SSB or was unconnected to current low levels of SSB and represented a decline in productivity associated with some extraneous environmental factor and was thus likely to persist, at least in the immediate future. The WG had undertaken a detailed analysis to explore the hypothesis that recruitment had declined as a result of a shrinking and warmer cold pool. Although recruitment variation was found to be associated with a measure indicative of the state of the cold pool, there was no evidence of a change in the latter measure that would explain the marked decrease in recruitment since 1990. A subjective examination during the review of the changes experienced over recent decades in a broader range of ecological and environmental variables failed to identify any specific factor that might explain the decline in recruitment.

While the weight of evidence presented at the review suggested that it was more likely than not (*i.e.,* a probability that the Panel subjectively assessed as roughly 60%) that the decline in recruitment of SNEMA yellowtail flounder was due to a productivity change associated with some unknown factor, the evidence was not sufficient to rule out the alternative hypothesis that the decline was the result of reduced SSB. The Panel therefore endorsed the WG's decision to present BRPs, stock status determinations, and projections for the two competing hypotheses. While it was concluded that, for both scenarios, overfishing was not occurring, the competing hypotheses produced very different conclusions regarding the overfished status of the stock. If the decline in recruitment since 1990 is the result of reduced SSB, it would be concluded that the stock is still overfished and has not yet been rebuilt, with projections indicating that it would be unlikely to rebuild by 2014 even in the absence of fishing. On the other hand, if, as the Panel concluded was more likely than not, the productivity of the stock has been reduced by some unknown factor, it would be concluded that the stock is rebuilt and not overfished. Note that, in this latter case, "rebuilding" is due to the reduction in the biomass reference point rather than increase in SSB.

The assessments produced for the GoM/GB Atlantic herring and SNEMA yellowtail flounder are of a high quality and provide the best scientific advice regarding the status of these two stocks that is currently available. The WGs are commended for their efforts in developing these assessments.

2. Background

2.1. Overview

A Stock Assessment Review Committee (SARC) meeting to review the 2012 benchmark stock assessments for Southern New England yellowtail flounder (*Pleuronectes ferrugineus*) and Atlantic herring (*Clupea harengus*) was held at the Northeast Fisheries Science Center, Woods Hole, Massachusetts, from 5-9 June, 2012. The SARC Review Panel for the 54th Stock Assessment Workshop (SAW 54) comprised, as chairman, Dr Robert O'Boyle (Beta Scientific Consulting, Canada, member of the New England Fishery Management Council's Scientific and Statistical Committee (SSC)), and, as panel members appointed by the Center for Independent Experts (CIE), Chris Francis (NIWA, NZ), Neil Klaer (CSIRO, AU), and Norman Hall (Murdoch Univ., AU) (Appendix 3). The agenda for the Review Workshop is presented in Annex 3 of Appendix 2.

The Statement of Work provided to Dr Norm Hall by the CIE is attached as Appendix 2. This required that, in addition to satisfying the requirement for SARC Panel members to participate in the review and conduct an independent peer review of each assessment, Review Panel members should assist the Review Chairman in preparing a SARC Summary Report of the review, and each should also prepare an independent CIE report of the assessments and the review process. This CIE report, which is prepared in accordance with the last of these requirements, describes my evaluation of the assessments and the review process.

Prior to the SARC Review Meeting, the stock assessment documents and other background documentation had been made available to Panel members. A list of these documents is presented in Appendix 1.

2.2. Terms of Reference

The terms of reference for the stock assessments of the Southern New England yellowtail flounder and Atlantic herring are presented in the Statement of Work (Appendix 2), together with the terms of reference for the SARC review of these assessments.

2.3. Panel membership

Details of the Panel Membership and of other key participants for the SARC review of the SAW 54 stock assessments for Southern New England yellowtail flounder and Atlantic herring are presented in Appendix 3. In particular, the SARC Review Panel members comprised:

- Robert O'Boyle, Panel Chair, member of NE FMC's SSC
- Chris Francis, CIE
- Neil Klaer, CIE
- Norman Hall, CIE

2.4. *Date and place*

The SARC met on 5-9 June, 2012, at the Northeast Fisheries Science Center, Woods Hole, Massachusetts, to review the benchmark stock assessments for Southern New England yellowtail flounder and Atlantic herring that had been produced for SAW 54.

2.5. *Acknowledgments*

Thanks are expressed to the various individuals who participated in the review meeting, and who contributed to the stock assessments, for making the review such an interesting and positive experience The WGs and, in particular, the presenters, J. Deroba and L. Alade, are to be commended for the quality of their stock assessments, and their very competent and professional responses to the Panel's queries and requests. Thanks are also extended to the NEFSC SAW Chairman, J. Weinberg, A. O'Brien, and P. Rago (NEFSC) for their assistance in providing access to review materials and ensuring the smooth running of the review meeting, and to the rapporteurs, T. Chute and J. Blaycock, who greatly assisted the Panel by recording the herring and yellowtail flounder discussions. The valuable insights, comments, and recommendations offered during the review meeting by R. O'Boyle, C. Francis, and N. Klaer are gratefully acknowledged.

3. Description of Reviewer's role in review activities

Prior to the review meeting, I familiarised myself with the background documentation, and the assessment and draft assessment summary reports for the two species that were the subject of the review (Appendix 1). Subsequently I attended and actively participated as a Review Panel member in the SARC meeting that was held at Woods Hole. At this meeting, the lead assessment scientists presented details of the data and the assessments, which I and the other Panel members reviewed and assessed. Together with other Panel members, I requested further details regarding specific aspects of the assessment that were of concern and considered and discussed the results of additional sensitivity runs that were requested during the meeting. I also participated in the Review Panel's discussions regarding the adequacy and soundness of the WG's responses to their various terms of reference, and whether the results of the assessments were of an appropriate scientific standard and thus acceptable as the basis for scientific advice for use in management. With other Panel Members, I contributed my suggestions of points to be considered when preparing the SARC Summary Report. Following the Review Meeting, I drafted those sections of the SARC Summary Report, for which I had been assigned responsibility, and offered comment on the resulting draft report. I then focused on preparing this document, *i.e.*, the CIE report describing my evaluation of the two stock assessments and the SARC review.

4. Summary of findings relevant to SARC review of the Terms of Reference of the stock assessments for SAW 54

In this section of the document, I have attempted to present my own assessment of each of the Terms of Reference for the assessments produced for SAW 54. Note however that, although the Statement of Work calls for an independent assessment by each CIE Panel Member, the review process itself and the process of preparing a SARC Summary Report that encompasses the different perspectives of the members of the Review Panel encourages convergence of views by Panel members and a focus on common issues. Note also that I have not attempted to paraphrase several sections of the SARC 54 Panel Summary Report, which I was responsible for drafting and which I have included in my CIE report.

*4.1 Atlantic herring (***Clupea harengus***).*

AH ToR 1. Estimate catch from all sources including landings and discards. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.

Was the ToR completed successfully?

Yes

Reason for acceptance/rejection

Catches from 1964 to 2011 for the Gulf of Maine/Georges Bank Atlantic herring stock complex by US fixed and mobile fishing gears and by the New Brunswick weir fishery were collated and reported. Discards were only included in catch estimates from 1996. As discards represented generally less than 1% of landings, the WG had considered that the lack of earlier historical data was unlikely to affect the results of the assessment. Age compositions (for ages 1 to 14) from 1964 to 2011 were reported for catches by US fixed and mobile fishing gears and (for ages 1 to 11+) from 1965 to 2011 for those from the New Brunswick weir fishery. Weights at age and proportions mature at age were reported. Plots showing the spatial distribution of herring landings per month from 2007 to 2010 were presented. Although the uncertainty of the catch estimates and biological data were not discussed explicitly in the section of the assessment report dealing with this ToR, uncertainty was incorporated in the assessment through iterative reweighting of the catch CV and estimation of effective sample size for the age compositions of the catches.

Strength of analysis

The Panel was advised in the review meeting that, although observer coverage was only about 5 – 15% in the early 2000s, recent observer coverage has been approximately 30%, with 100% coverage in closed area 1 and 80% coverage on Georges Bank. Thus, overall, coverage has been sufficient to conclude that discards have generally been low.

The fact that strong and weak year classes are evident in the age composition data indicates that the age composition data contain information on year class strength that is likely to be of value to the assessment.

Weakness of analysis

No detail is provided in the Assessment Report of the sampling design and methods used to collect biological samples, nor the methods used to analyse and expand the resulting data to produce the tables of age compositions of annual catches, catch weights at age, spawning stock biomass at age, and proportions mature at age. It was thus not possible to assess how well the reported data were likely to reflect the true age compositions, weights at age, etc.

The uncertainty of the estimates of catch, catch at age, weight at age and maturity data was not discussed in the section of the assessment report dealing with this ToR.

The catch history that has been developed relies on the adequacy of the assumption that the stock complex and its fishery are constrained spatially to lie within the geographic boundaries that were selected by the WG, and that there is no movement of fish across those boundaries or interchange between stocks/fisheries. There would be value in developing alternative catch histories that allow for uncertainty in the northern boundary of the stock, and possible mixing of fish between the Gulf of Main/Georges Bank and Scotian shelf stocks and the effect of such mixing on catches. By exploring the sensitivity of the assessment to these alternative catch histories, insight would be gained on the uncertainty associated with catches, discards, and the spatial distribution of the stock complex and the catches taken from that complex.

Were conclusions and recommendations acceptable?

The decisions of the WG to use the catch data for the fixed and mobile fishing gears from 1965 to 2011 in the assessment and to pool data for older fish into an 8+ age class were acceptable.

Reason for acceptance/rejection

The similarity of the fishing gears and of age composition data for the New Brunswick weir and US fixed gear fisheries supports the decision to combine the data for these fisheries and treat them as a single time series of catches. The length frequencies and age compositions of the catches from the fixed and mobile gears differ markedly, supporting the decision to treat the data for these two gear types as separate time series of catches. The paucity of data for fish older than age 8 supports the decision to pool these data into an 8+ age class.

Does work provide a scientifically credible basis for fishery management advice?

The catch and age composition data provide a sound basis for their use in developing assessment models aimed at providing fishery management advice.

AH ToR 2. Present the survey data being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, larval surveys, age-length data, predator consumption rates, etc.). Investigate the utility of commercial LPUE as a measure of relative abundance, and characterize the uncertainty and any bias in these sources of data.

Was the ToR completed successfully?

Yes

Reason for acceptance/rejection

The WG considered and presented abundance and, where available, age composition data for each of the surveys that was considered to be a candidate for use in the assessment, *i.e.*, the NMFS spring and fall bottom trawl surveys, the NMFS winter survey, the NMFS summer shrimp survey, the ichthyoplankton larval surveys, the Massachusetts Department of Marine Fisheries (MA DMF) spring and fall bottom trawl surveys, and the Maine/New Hampshire spring and fall bottom trawl surveys. Predator consumption rates were considered under ToR 6. The potential use of commercial LPUE was considered and discussed by the WG in its assessment report. Uncertainty of abundance data was characterised by the CV of the survey abundance estimates. Although no details of sample sizes were presented for the age composition data, estimates of effective sample size were determined through the iterative approach described by the WG under ToR 5. The MA DMF and Maine/New Hampshire surveys were eliminated from use in the assessment because of inadequate spatial coverage of the stock complex, while the NMFS winter survey was eliminated by the WG due to inconsistent spatial coverage and lack of fit. Because the relationships of the larval index to both its parent spawning stock biomass and its subsequent age-1 recruitment were unclear, the WG eliminated this index. Factors that affected the consistency of the NMFS spring and fall surveys, such as change of trawl boards or change of survey vessel, were also considered and dealt with by breaking the time series at 1985, in the case of the change in trawl boards, or through adjustment by calibration factors in the case of survey vessel change.

Strength of analysis

The long time series of spring and fall survey data, collected using a sampling protocol that provides consistent spatial coverage, provide valuable information on trends in abundance and age composition of Atlantic herring.

The decision not to apply age-length keys derived from commercial catch data to survey length data collected prior to 1987 is strongly endorsed as such application is inappropriate and would introduce error.

Weakness of analysis

Although unavoidable, the breaks in consistency of the nets, trawl gear and vessels used in the spring and fall surveys have reduced the value of these two long time series.

While the pattern of residuals for the annual spring and fall survey indices supports a decision to break the two time series in 1985, the trend in the residuals of an earlier model that did not break each time series into two sections suggests that the transition occurred over a number of years, and was not entirely consistent in timing with the change in trawl doors. While the decision to break the time series at 1985 is endorsed, consideration should be given to identifying alternative hypotheses that might explain the trend in the residuals from the earlier model.

The trend with length of the length-specific calibration factors used to convert data collected by the FSV Henry B. Bigelow for the fall survey to FRV Albatross IV equivalent indices is unusual and a process to justify such complexity should be identified.

Uncertainty associated with the calibration factors used to convert data collected by the FRV Delaware II and FSV Henry B. Bigelow to FRV Albatross IV equivalent indices is not carried through to the results of the assessment model.

Were conclusions and recommendations acceptable?

Yes

Reason for acceptance/rejection

The decision to use the NMFS spring and fall surveys, and the NMFS summer shrimp survey, in the assessment was justified as each survey was considered to provide good spatial coverage of the stock complex, and thus survey indices and age compositions were likely to be representative of the stock.

Rejection of the state surveys was justified as inadequate coverage of the Atlantic herring stock complex would lead to a non-representative sample of the stock complex with abundance and age composition estimates reflecting only the abundance and age composition of that portion of the stock complex available within the area covered by the state surveys. For the state survey data to serve as a useful indicator, it would be required that the proportion of each age class available within the surveyed area is constant among years, although possibly varying among age classes. Further exploration should be undertaken to assess the validity of such an assumption before any decision is made to incorporate data from the state surveys as indices of abundance in the assessment.

The decision to reject the NMFS winter survey as an index of abundance is justified as inconsistent spatial coverage will be reflected in the indices of abundance that are calculated. The resulting indices of abundance would not reflect the trend in abundance of the stock complex.

The decision to reject the larval survey index was justified because of the WG's uncertainty that this index reflected the abundance of either the parent spawning biomass or the subsequent age-1 recruitment.

Rejection of commercial LPUE as an index of abundance of Atlantic herring is justified, given the changes in spatial distribution of fishing that have resulted from regulation changes, the potential for hyperstability resulting from the use of sonar to locate and track schools, and the difficulty in identifying fishing effort directed towards Atlantic herring due to target switching during trips.

Does work provide a scientifically credible basis for fishery management advice?

Abundance and age composition indices calculated from the NMFS spring and fall surveys, and the NMFS summer shrimp survey, provide a scientifically credible basis for fishery management advice.

AH ToR 3. Evaluate the utility of the NEFSC fall acoustic survey to the stock assessment of herring. Consider degree of spatial and temporal overlap between the survey and the stock. Compare acoustic survey results with measures derived from bottom trawl surveys.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG presented details of the NEFSC fall acoustic survey and considered the trend in the resulting estimates of abundance of Atlantic herring, which exhibited an abrupt and marked decline from values observed till 2001 to those observed in subsequent years. Plots of survey tracks relative to the spatial extent of the areas of the Georges Bank used in the 1999 and 2012 assessments, which could be compared with plots of the spatial distributions of monthly catches, were presented in the Assessment Report. The trends in the estimates of abundance calculated using the acoustic survey data were compared with those derived from other bottom trawl surveys.

Weakness of analysis

When comparing the estimates for 2006 from the NEFSC fall acoustic survey with those from the long range sonar (OAWRS) study, it is stated in the assessment report that "All approaches were consistent to within 20% or less". This statement appears inconsistent, however, with the values shown in Table A3-4 of the Assessment Report.

Were conclusions and recommendations acceptable?

The decision to exclude the acoustic survey index from the assessment was acceptable.

Reason for acceptance/rejection

The marked decline in the acoustic index from the values observed between 1999 and 2001 to those observed in subsequent years was inconsistent with the trends in other survey indices and fishery monitoring data. Although several hypotheses had been proposed to explain the decline, there was insufficient evidence in support of any specific hypothesis to justify its incorporation in the model. The suggestion that the decrease could be explained by a mis-match between the timing of the survey and the time at which spawning occurred was discounted by the results from the larval study, which back-calculated spawning time.

Other hypotheses considered were differences in fish behaviour and vertical distribution during the different years of the acoustic survey, but the factor(s) involved has/have yet to be identified. It should be noted, however, that the acoustic survey extends over a limited spatial area and indices are thus representative of only a portion of the stock, whereas the other survey indices are representative of the entire stock complex.

Does work provide a scientifically credible basis for fishery management advice?

The decision to exclude the acoustic survey index of abundance was appropriate, as its inclusion would have been likely to have led to biased fishery management advice.

AH ToR 4. Evaluate the validity of the current stock definition, and determine whether it should be changed. Take into account what is known about migration among stock areas.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG collated data on the spatial distribution of Atlantic herring derived from fishery-independent survey and ichthyoplankton data, geographic variation in biochemistry, growth, and morphology, and information on movements and migration derived from ichthyoplankton dispersion data and tagging studies (both within and outside the Gulf of Maine/Georges Bank region). It then re-examined those data to determine whether the current stock definitions should be changed.

Strength of analysis

The re-assessment was based on a comprehensive review of available data.

Weakness of analysis

The decision to maintain the current stock definition, *i.e.*, to treat Atlantic herring in the Gulf of Maine/Georges Bank as a stock complex, was based on the fact it is currently not possible to distinguish from which of the different component sub-stocks within this region the fish in commercial and survey catches are drawn. It was also based on the assumption that fish from the Scotian Shelf stock and those from the Gulf of Maine/Georges Bank stock complex remain separate. The WG "acknowledged some degree of mixing of Scotian shelf stocks with U.S. stocks", suggesting that there would be value in considering the sensitivity of the results from the base model for the 2012 stock assessment to an alternative stock structure with some level of mixing of fish from the Scotian shelf in catches and survey data, *e.g.*, through use of alternative catch histories or movement rates from tagging studies. Currently, there is no assessment of the uncertainty associated with the stock structure that has been adopted by the WG.

The WG reported in its Assessment Report that several studies, which had examined morphometric and otolith variation, had successfully identified the stock of origin with considerable (70 to 88%) accuracy. If further research demonstrates that this approach is reliable, it may offer the opportunity to classify sampled fish from commercial and survey catches according to their stock of origin.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The decision to treat the Gulf of Maine and Georges Bank stocks as a single stock complex is justified by the fact that the stock of origin of fish in commercial and survey catches cannot be identified.

Does work provide a scientifically credible basis for fishery management advice?

Yes. While the stock definition is considered appropriate for current use in determining management advice, the potential exists that, even though the stock complex may be assessed as not overfished with overfishing not occurring, a less-productive sub-stock within this complex may be overfished or being subjected to overfishing.

AH ToR 5. Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from AH TOR-6), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The ASAP model used for the 2009 TRAC assessment was updated by the WG with data through 2011. Because of the retrospective problems encountered with the 2009 assessment, which remained present with the extended time series, all data inputs and model settings were reviewed. While the 1999 ASAP model was fitted to pooled catch data and assumed a selectivity of 1, the revised base model was fitted to the time series of catches for fixed and mobile gears and estimated selectivity. Selectivity was also estimated for the survey data. In contrast to the 1999 assessment, age compositions for survey data prior to 1987, for which no age data had been collected, were not estimated using age-length keys derived from commercial data as had been done in the previous assessment. Age 8 was treated as a plus group. The key modification to the assessment was the replacement of the assumption that natural mortality M was equal to 0.2 by the assumption that, for each year, the average level of natural mortality across all ages was 0.3, and that the relative level of natural mortality at age declined with average weight at age in accordance with relative levels of natural mortality predicted by the equation derived by Lorenzen (1996) for ocean fish. The value of 0.3 was derived from a maximum age for Atlantic herring of 14 years, using Hoenig's (1983) relationship between mortality and maximum age for fish. The resulting estimates for natural mortality for each age were smoothed across years using a General Additive Model (GAM). Age-dependent mortalities from 1996 to 2011 were increased by 50% to reflect the increased consumption of Atlantic herring that had occurred in those years. This increase largely resolved the retrospective pattern in SSB that had been present.

Estimates of annual fishing mortality, age-1 recruitment, spawning stock biomass and total biomass calculated using the base model were reported by the WG in their Assessment Report. The posterior probability distributions of SSBs and Fs were determined using MCMC and time series plots of these variables, with 80% confidence limits, were plotted. Similar results for age-1 recruitment and total biomass were not presented in the Assessment Report, however. Sensitivities of SSB and F to the inclusion/exclusion of the winter, larval and acoustic surveys and to use of F=0.2 for all ages and years were explored. The WG also presented a plot showing the sensitivity of the estimates of SSB, F, and age-1 recruitment, but not total biomass, to an ASAP run using fish consumption as a fleet and a run without the 50% increase in natural mortality from 1996 to 2011. The sensitivity of the values of F and SSB relative to respective reference points for varying values of steepness of the Beverton and Holt stock-recruitment relationship was also examined.

Retrospective patterns for SSB, F, and age-1 recruitment were explored for the base run, and the time series of estimated SSBs and Fs from the 1995, 2005, and 2009 assessment models were compared with those from the base run. No comparison with previous projections was made.

Strength of analysis

Use of the ASAP model, which has been well tested, was appropriate for this assessment. It was also pleasing to note that other model structures, such as SS3, are being explored.

The uncertainty in the assessment was explored using MCMC and alternative runs to assess sensitivity to model assumptions.

Examination of the likelihood profile demonstrated that the steepness parameter of the stock-recruitment relationship was very imprecise, with a 95% confidence interval ranging from less than 0.35 to 0.85 and with associated biological reference points and estimates of SSB and fishing mortality that varied greatly with the changes in steepness. Despite this, the conclusions regarding stock status remained unchanged from the findings obtained from the base model over the range of steepness values except in the case of an atypically low steepness of 0.35, where the results indicated that overfishing was occurring.

Weakness of analysis

The use of the Lorenzen (1996) model to introduce an age-dependent mortality is not justified as it increases model complexity and, in the case of the Atlantic herring model, is based on an assumption rather than support from observed data. Furthermore, advice was provided at the review meeting that the model using Lorenzen-based age-dependent mortalities failed to improve the fit relative to a model that assumed constant M at age.

The Panel recommended that effective sample sizes for age compositions should be calculated using iterative reweighting based on mean age, and possibly reflecting relative magnitude of initial sample sizes.

It is recommended that a more appropriate CV be imposed on recruitment deviations. Mertz and Myers (1996) report that, for most fish, the value of $\sigma_{\log_e R}$ falls within the range from 0.3 to 1. The base model imposed a minimal constraint on the deviations, *i.e.*, CV=1, which may have influenced the 2009 age-1 recruitment estimate, which was estimated to be approximately twice as great as the maximum of recruitment estimates for earlier years and which, as a consequence, influenced estimates of recent biomass.

The substantial uncertainties in the calibration factors employed to adjust the survey data for the effect of changes in survey vessel could not be carried through to the estimates produced by the model.

The effect of increased or decreased weighting for different survey or age composition data on stock assessment results was not examined, and thus it was not possible to assess whether the different survey indices exhibited tension in the support they exhibited for different values of key model parameters.

The extent to which survey catchability was estimated to increase when the trawl doors were changed in 1985 was very different in the fall and spring surveys, *i.e.*, 2.64 for the spring survey *cf.* 13.6 for the fall survey. This, combined with the fact that the residuals for the spring and fall indices for an earlier model that assumed constant catchability for each time series exhibited a trend suggesting transition over a number of years and timing that was not entirely consistent with the trawl door change, suggests that another factor may have been involved. This deserves further investigation.

Were conclusions and recommendations acceptable?

The Panel accepted the base case, which was based on M=0.3, with a 50% increase for years after 1995, as the most plausible model for provision of management advice.

Reason for acceptance/rejection

The temporal trend in consumption data for Atlantic herring justifies the assumption that M would have increased in the later years of the time series. While inclusion of the consumption data in an alternative model was explored, the WG concluded that "the inter-annual variation of the fish predator consumption estimates was not well understood and was beyond what would be expected from a relatively constant predator fleet", and such a model was therefore eliminated from further consideration. Nevertheless, the average of the consumption estimates was clearly greater from about 1996, and thus, although unusual for an assessment model, an assumption that M increased in this later period is justified.

The Panel examined the results obtained during the review meeting by fitting the base model but using alternative values of percentage increase in post-1995 natural mortality. Estimates of likelihood, degree of retrospective pattern (Mohn's Rho), and extent to which the average level of implied consumption in the post-1995 period matched the estimates of consumption were compared. From this, the Panel concluded that the 50% increase in post-1995 mortality assumed in the base model was appropriate and plausible, producing post-1995 consumption estimates that were realistic and retrospective patterns that were acceptable.

The Panel was concerned that the estimate of recruitment for the large 2008 year class might have been affected by the relatively large CV that had been imposed on recruitment deviations in the base case model. However, although an exploratory run using a reduced CV lessened the estimated strength of the 2009 age-1 recruits, the resulting recruitment estimate for the 2008 year class remained the highest in the time series.

Does work provide a scientifically credible basis for fishery management advice?

Yes. The base case model produced by the WG provides a credible basis for use in developing fishery management advice. Continued high levels of consumption of Atlantic herring by fish predators is expected in the immediate future, but monitoring should be undertaken to assess whether this persists in the longer term.

AH ToR 6. Consider the implications of consumption of herring, at various life stages, for use in estimating herring natural mortality rate (M) and to inform the herring stock-recruitment relationship. Characterize the uncertainty of the consumption estimates. If possible integrate the results into the stock assessment.

Was the ToR completed successfully?

The implications of consumption of herring, at various life stages, for use in estimating herring natural mortality rate (M) and to inform the herring stock-recruitment relationship were considered and the results were used to inform the stock assessment. The uncertainty of the estimates of consumption by fish predation was not characterized.

Reason for acceptance/rejection

The WG undertook a comprehensive examination of the data on predator abundance and associated consumption data. The predators, which were considered, included fish, two highly migratory species of fish, marine mammals and sea birds. Estimates of consumption were calculated by combining estimates of predator abundance with estimates of total consumption per predator and dietary composition. Although estimates of consumption of Atlantic herring were produced, the uncertainty of these consumption estimates was not addressed or carried into the assessment. Estimates of consumption were included in an exploratory run of the assessment model but did not improve model fit. Estimates of consumption were used, however, to inform the decision to increase post-1995 natural mortality in the base assessment model.

Strength of analysis

The comprehensive data that are available from stomach analyses for fish from the Gulf of Maine/Georges Bank region provides a strong basis for developing estimates of the consumption of Atlantic herring by the various fish predators. This, coupled with estimates of predator abundance from long time series of fishery-independent survey data, has allowed the construction of a time series of annual consumption of Atlantic herring, which, given the uncertainties involved, appears reasonably sound.

It was noted by the Panel that most of the annual signal in the consumption of herring was related to stomach contents rather than predator abundance.

Weakness of analysis

Abundance estimates of some predator species were calculated from swept area calculations, rather than assessment models.

The later peak in the time series of annual consumption estimates was driven by the very high abundances of two individual predator species, but similar high levels of abundance for those species were not present in adjacent years. Greater inter-annual consistency might have been expected.

The uncertainty of the consumption estimates needs to be assessed.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

It is unusual to consider annual consumption in stock assessment as, in most cases, a time series of estimates of this variable is not available. In the case of the Gulf of Maine/Georges Bank Atlantic herring, however, the wealth of available annual information from stomach contents from a wide range of herring predators (fish, mammals, other), and annual abundance estimates of those predators, has afforded the opportunity to calculate a time series of the annual consumption for this stock complex. Despite high uncertainty in the data, the consumption estimates provide strong evidence that consumption of herring has increased since 1995. It was noted, however, that the consumption estimates, which have been calculated, are likely to be an underestimate of total consumption and that inclusion of estimates of consumption in an exploratory run of the assessment model did not improve model fit.

Advice received by the Panel during the review meeting indicated that, given the current trends in abundance of predator populations, consumption of Atlantic herring is likely to remain high.

Does work provide a scientifically credible basis for fishery management advice?

Yes, but the uncertainty of consumption estimates needs to be recognised.

AH ToR 7. State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

Existing stock status definitions for "overfished" and "overfishing", which were based on the fit of a Fox surplus production model, were reported in the Assessment Report. Updated MSY reference points, which were based on the Beverton and Holt (BH) stock-recruitment relationship fitted internally in the base model, were derived. Values of these reference points, together with 80% confidence limits determined from an MCMC analysis, were reported. The WG commented on the scientific adequacy of the existing and new reference points and concluded that the new reference points were an improvement on the existing reference points.

Strength of analysis

The analysis was sound. Uncertainty in the reference points associated with age-specific M, the percentage increase in M since 1995, and steepness of the Beverton and Holt stock-recruitment relationship, was explored.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The Panel noted that the existing biological reference points were derived using an age-aggregated rather than age-structured model, and based on an analysis on catch data aggregated over gear type rather than an analysis that employed catch by mobile or fixed gear. The Panel endorsed the redefinition of the biological reference points to those derived using the base model from this assessment.

It was recognised that, as these biological reference points were based on conditions prevailing in 2011, the WG assumed a continued level of high natural mortality due to consumption by predators. This assumption was considered appropriate for the immediate future, *i.e.*, for the next three to five years, but there was uncertainty as to whether such high mortality would continue to prevail in the longer term. Continued monitoring of predation was recommended.

Does work provide a scientifically credible basis for fishery management advice?

Yes. The updated estimates of the BRPs, *i.e.*, Fmsy, SSBmsy, MSY, and the value of ½MSY used as the reference level when assessing whether the stock is overfished, which were determined using the Beverton and Holt stock-recruitment relationship fitted internally in the new base model, provide a scientifically credible basis for fishery management advice.

AH ToR 8. Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model, should one be developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).

a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.

b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from AH TOR-7).

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The existing model was updated by the WG with data through 2011 and stock status was evaluated with respect to the existing reference points. Stock status was also evaluated using the estimates and new MSY reference points determined using the base model for the 2012 assessment.

Strength of analysis

The status of the stock was determined for a range of both age-specific and post-1995 values of M, and for a range of values of the steepness parameter of the stock-recruitment relationship. Apart from the unlikely case in which steepness was assumed to be 0.35, the conclusion was that the stock was not overfished and overfishing was not occurring.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The Panel agreed that, while there is uncertainty in the values of the reference points and estimates of SSB and F, the conclusions that the stock is not overfished and that overfishing is not occurring appeared robust to that uncertainty. The conclusion using the previous model was consistent with that from the new base model and new reference points.

Does work provide a scientifically credible basis for fishery management advice?

Yes. The assessment indicates that the stock is not overfished and that overfishing is not occurring.

AH ToR 9. Using simulation/estimation methods, evaluate consequences of alternative harvest policies in light of uncertainties in model formulation, presence of retrospective patterns, and incomplete information on magnitude and variability in M.

Was the ToR completed successfully?

No, although some initial work is underway.

Reason for acceptance/rejection

Until now, it appears that the focus of the WG has been on developing the ASAP model and exploring its uncertainties. Although research projects have been initiated to explore some aspects of model fitting and assessment of stock status, *e.g.*, the consequences of ignoring retrospective patterns, no studies to undertake a formal management strategy evaluation have yet been initiated.

The Panel recommended that development of a management strategy evaluation framework for the Atlantic herring should be pursued, but suggested that, as a first step, this would require the engagement of fishery managers and stakeholders to consider the specification of formal harvest control rules and identify the alternative harvest policies that should be explored. Management Strategy Evaluation would assist in exploration of strategies to achieve management objectives under the largest uncertainties that exist, *e.g.*, whether high values of natural mortality will persist.

AH ToR 10. Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).

b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The base model was employed by the WG to project the likely outcomes in 2013, 2014, and 2015 of employing alternative fishing mortalities or catches. Probabilities of exceeding threshold BRPs for F or falling below threshold BRPs for SSB were computed, and the implications of uncertainty in steepness and natural mortality were explored. Factors such as the unknown contribution of fish from the Scotian shelf and the possibility that the strength of the 2008 year class is overestimated, which would affect the stock's vulnerability to becoming overfished, were discussed.

Strength of analysis

Assessment uncertainties were well described.

Drawing numbers-at-age in 2012 from the results of the MCMC simulations of the base ASAP run ensures that uncertainty of the initial age composition at the start of the projection period is well characterised.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The WG has explored the effect on projections of a range of assumptions and the most important sources of uncertainty, *i.e.*, age-specific M, percentage increase in post-1995 M, steepness of the stock-recruitment relationship and size of the 2008 year class.

Vulnerability of the stock to various factors, such as (1) whether the high natural mortality experienced in recent years is likely to continue over the next three to five years, (2) whether the size of the 2008 year class is as large as has been estimated, and (3) the fact that the assessment is based on data for a stock complex rather than a single stock and that fish from the Scotian shelf make an unknown contribution to catches, were considered.

Does work provide a scientifically credible basis for fishery management advice?

Yes. Projections made using the base model for a range of alternative harvest scenarios provide a credible basis for fishery management advice.

AH ToR 11. For any research recommendations listed in recent peer reviewed assessment and review panel reports, review, evaluate and report on the status of those research recommendations. Identify new research recommendations.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG determined that no research recommendations had been proposed at previous assessments. Research recommendations arising from the current assessment were reported.

Were conclusions and recommendations acceptable?

Yes.

Reason for acceptance/rejection

The list of recommendations proposed by the WG in the Assessment Report needs to be sorted into priority order. Consideration should be given to modifying recommendation f to "Determine the depth distribution of Atlantic herring and the factors that influence that distribution". Recommendation g appears too vague and requires clarification. The recommendations to collect additional data and to improve the discrimination of stocks and sub-stocks, and to monitor the abundance of the various sub-stocks are endorsed, as modelling the Gulf of Maine/Georges Bank Atlantic herring as a stock complex masks the possibility that a less-productive sub-stock might be overfished or experiencing overfishing. The recommendations associated with improving the estimation of the consumption of Atlantic herring are also endorsed as they offer the potential that future models might explicitly incorporate consumption of herring as an explanatory variable affecting natural mortality, thereby allowing the variability and uncertainty associated with this component of natural mortality to be carried through to estimates of BRPs and determination of stock status. Other recommendations related to improvements to data and methods for stock assessment.

The Review Panel identified a number of research areas that should also be considered. These are listed below.

Data

- Develop alternative catch histories to allow for interchange of fish from the Gulf of Maine/Georges Bank stock complex with those from the Scotian shelf stock, and use these to explore the uncertainty associated with the boundary of the Gulf of Maine/Georges Bank stock area.
- Use the alternative catch histories developed in the previous research item to explore whether catch under-estimation (*e.g.,* inclusion of Scotian shelf catch) can contribute to the reduction in the retrospective pattern and contribute to or explain the need for increased M.

Model

- Explore the effect of imposing a penalty to introduce greater similarity in the relative increase in catchability experienced by the spring and fall survey data in response to the change in trawl doors.

Management Strategy Evaluation (MSE)

- Jointly with managers and stakeholders, identify formal harvest control rules and harvest strategies to be explored, and major uncertainties to be considered, and based on this information, undertake a management strategy evaluation.

*4.2 Southern New England yellowtail flounder (***Pleuronectes ferrugineus***)*

YF ToR 1. Estimate landings and discards by gear type and where possible by fleet, from all sources. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.

Was the ToR completed successfully?

Yes.

Reason for acceptance/rejection

The WG collated and reported landings and estimated discards of SNEMA yellowtail flounder between 1935 and 2011. Of these data, only the time series of total catches (landings + dead discards) from 1973 to 2011 was employed in the assessment. For 1994 to 2011, the Assessment Report provided estimates of landings (and CVs) by commercial fishers employing trawl (88 to 99% of total landings), scallop dredge, gillnet, and other/unknown fishing gear and fishing in the area over which the stock is distributed. Discards (and CVs) were also collated and reported (by gear type) for the same period. Catch at age of commercial landings between 1994 and 2011 was calculated using a revised length-weight relationship. Plots of the time series of landings reported from catches from different statistical areas were included in the Assessment Report.

Strength of analysis

The data reported for the 2012 assessment reflect a marked revision of data and assumptions from the previous assessment in 2008.

The precision of estimates of discards has been improved by use of spatial stratification.

The revised length-weight relationships have improved estimates of weight at age, and, through their effect on conversion of catch weights to numbers at age, have led to improved estimates of age compositions.

An improved estimate of discard mortality was determined using results from a Reflex Action Mortality Predictor (RAMP) study. The decision to use 90% as the

5. Conclusions and recommendations

The SARC 54 Review Panel examined the 2012 stock assessments developed for the Southern New England yellowtail flounder (*Pleuronectes ferrugineus*) and Atlantic herring (*Clupea harengus*). After considering the information relating to stock structure, the data that were available, and the details of the assessment for each species, and discussing areas of concern, the Panel accepted the base model that had been proposed by the WG for each assessment. The accepted base models, which had been developed using the statistical catch-at-age ASAP model, differed from those which had been applied in the previous assessments, *i.e.*, the 2009 TRAC assessment for Atlantic herring and the 2008 GARM assessment for yellowtail flounder.

The WG's decision to maintain the current stock definition for the Gulf of Maine/Georges Bank Atlantic herring was endorsed by the Panel, which recognised that this "stock" was a complex that comprised at least three "sub-stocks". In the case of this stock, the assumption of a constant natural mortality of M=0.2 $year^{-1}$ in the ASAP model, which had been used in the previous 2009 assessment, had been replaced by an assumption that natural mortality declined with weight at age in a pattern similar to that described by Lorenzen (1996) and thus varied with both age and year, that the average value of natural mortality over age in the years up to 1995 was 0.3 $year^{-1}$, and that, since 1995, the average value of natural mortality over age had increased by 50%. Rather than fitting the model to total catch, as was done in the 2009 assessment, catches of fixed (US fixed gear and New Brunswick weir fishery) and mobile gears were used in the 2012 assessment. Maturity at age, which was constant in the 2009 assessment, varied with age in the new assessment.

The Panel was advised that, in the case of Atlantic herring, introduction of age and time dependent natural mortality, based on Lorenzen (1996), failed to improve the overall likelihood of the model relative to a model that employed a constant mortality over age and time. On this basis, the added complexity of the age and time dependent mortality is not justified. However, its inclusion is not likely to affect the results. A more important issue is whether or not natural mortality increased by 50% since 1995. This unusual increase in natural mortality was justified by increased annual consumption by predators of Atlantic herring, and by the fact that inclusion of the increased natural mortality reduced retrospective patterns. The very considerable body of data relating to dietary composition and predator abundance suggests that the trends in consumption are real, and that it is appropriate to increase natural mortality in the latter portion of the time series. Accordingly, after considering other results, diagnostic outputs, and output from various sensitivity runs, the Panel accepted the base ASAP assessment model for Atlantic herring as an appropriate model for developing management advice. While there is some uncertainty regarding the extent to which natural mortality increased, examination of the results obtained by applying a range of alternative percentage increases suggested that the value of 50% was plausible. It should be recognised, however, that the uncertainty associated with this percentage is not carried through into the uncertainty associated with estimates of spawning stock biomass (SSB), fishing mortality (F), or biological reference points (BRPs), or into projections.

As it is very likely that high levels of consumption of Atlantic herring will persist in the immediate future, the decision by the WG to employ the increased level of natural mortality when calculating BRPs was endorsed by the Panel. Likewise, the Panel concurred with the finding that the GoM/GB Atlantic herring stock complex was not overfished and that overfishing was not occurring. Exploration of the sensitivity of this finding to a range of values of M and post-1995 increase in M, and to a range of values of steepness of the BH stock-recruitment relationship employed in the ASAP model for the Atlantic herring demonstrated that this finding was robust to such uncertainty.

The major uncertainties of the 2012 assessment for GoM/GB Atlantic herring are the extent to which natural mortality increased since 1995, whether the high level of consumption by predators is likely to persist in the longer term, the extent to which mixing of fish from different stocks across the boundaries of the Gulf of Maine/Georges Bank region has influenced the assessment, whether any of the various sub-stocks in the Gulf of Maine/Georges Bank stock complex are overfished or experiencing overfishing, and whether the abundance of the 2008 year class, which is currently estimated to be twice as great as the largest year class previously encountered (1994) and which is projected to make a considerable contribution to future catches, has been over-estimated.

The Panel endorsed the WG's conclusion that yellowtail flounder in the southern New England/Mid-Atlantic (SNEMA) region should be considered as a separate stock for management purposes. For this stock, the assumption of a constant natural mortality of 0.2 $year^{-1}$ in the VPA model, which had been used in the previous 2008 assessment, was replaced by an assumption that natural mortality declined with weight at age in a pattern similar to that described by Lorenzen (1996) and thus varied with both age and year, and that the average value of natural mortality over age was 0.3 $year^{-1}$. The data series used had also been revised considerably since the previous assessment. Discards for the 2012 assessment were estimated using spatial and temporal stratification of observer data, estimates of discard mortality were revised, and, from 1994, revised weight-length relationships were used to convert catches to numbers at age. Estimates of weights and maturity at age were also updated.

The WG is commended for the bridging analyses that they undertook to demonstrate the effect of each step of the transition from the previous VPA model for yellowtail flounder to the new ASAP model with its revised data.

The Panel did not support the WG's conclusion that estimates of abundance from daylight survey trawls were sufficiently similar to those from trawls at night that an aggregate index could be used and advised that improved precision would be gained by taking the time of day, *i.e.*, day or night, into account. The assumption in the new model that natural mortality is age and time dependent, and follows a pattern such as that described by Lorenzen (1996) introduces additional complexity to the model for yellowtail flounder. A simpler model that assumes constant natural mortality is to be preferred unless it can be demonstrated that the more complex model improves the likelihood significantly. Data weightings used in fitting the model, *e.g.*, doubling the larval CVs and adding 0.1 to survey CVs, appeared arbitrary.

After considering the results of the base assessment model and its diagnostic outputs, and the results and diagnostic output from various sensitivity runs, the Panel accepted the base ASAP assessment model for SNEMA yellowtail flounder as an appropriate model for developing management advice. A key result from this base model, however, was the marked decline in recruitment that had occurred since 1990. The WG proposed two scenarios to account for this. Firstly, the WG proposed a "two-stanza recruitment scenario" that assumed that recruitment is a function of SSB, with high but very variable recruitment at larger values of SSB and very low recruitment when SSB is less than about 4,300 mt. The WG's second proposal was a "recent recruitment scenario", which assumed that, since 1990, annual recruitment is low because of environmental factors that are unrelated to SSB. The Panel agreed that, given the pattern of recruitment predicted by the base model, these two recruitment scenarios were appropriate alternatives. The Panel also recommended that, in the future, consideration might also be given to exploring whether a BH model fitted internally within the ASAP model, might provide a further alternative. By allowing for a change in the values of the parameters of the stock-recruitment relationship in 1990, such a model could be used to explore the extent to which steepness or asymptotic recruitment had changed.

The Panel accepted the WG's finding that variables indicative of the size and temperature of the cold pool influenced the strength of recruitment of yellowtail flounder, but concluded that these variables provided no explanation of the step-like decline in recruitment in 1990. Likewise, data relating to broader ecosystem and environmental change in the region provided no explanation for the sudden decline in recruitment.

Based on the weight of the evidence that had been presented, the Panel concluded that, while it was not possible to rule out either of the two recruitment scenarios, it was more likely than not that productivity of the yellowtail flounder stock had declined since 1990, *i.e.*, there was a qualitatively-assessed probability of about 60% that current recruitment will follow the recent recruitment scenario rather than the two-stanza scenario. Note that the value of 60% is based on the Panel's subjective opinion, and has no quantitative basis. Recognising that the evidence was insufficient to rule out either scenario, however, the Panel agreed with the WG that both recruitment scenarios should be considered when evaluating reference points, determining stock status, and considering the results of short term projections.

Results from the base assessment model for SNEMA yellowtail flounder indicate that, for either recruitment scenario, overfishing is not occurring. If recruitment follows the two-stanza scenario, however, the stock is not rebuilt, is overfished, and is unlikely to rebuild by 2014 even in the absence of fishing. On the other hand, if the productivity of the stock has been reduced by some unknown factor and, as the Panel considers is more likely than not, recruitment follows the recent recruitment scenario, the stock is rebuilt and not overfished. Note that the fact that the stock is considered to be rebuilt in this latter case is due to the reduction in the biomass reference point rather than increase in SSB.

Key uncertainties in the case of the SNEMA yellowtail flounder assessment include the identity of the factor or factors that have led to and would explain the decline in recruitment since 1990, the question of whether low recruitment is likely to persist, the possibility of error in catches and discards (which could be explored through development of alternative catch histories), and the possibility that the SNEMA yellowtail flounder stock is not entirely closed to interchange of fish or larvae across the boundary of the SNEMA stock area.

The unusual features of the stock assessments undertaken for SWA 54, *i.e.*, that, since 1995, natural mortality of Atlantic herring had increased markedly, and that, since 1990, there had been a decline in the recruitment of yellowtail flounder, highlight the fact that the assumptions of a constant level of natural mortality and a constant stock-recruitment that are typically used in stock assessment models are probably over-simplistic. With global climate change, it would not be unexpected that natural mortality, growth, weight at age, maturity at age, and stock-recruitment of other stocks are also experiencing change. In most cases, however, the change would be expected to be of a more gradual nature than the changes that appear to have occurred in the cases of natural mortality of GoM/GB Atlantic herring or recruitment of SNEMA yellowtail flounder.

Rather than constructing separate models to describe alternative scenarios, as was done in the case of yellowtail flounder, there would be value in developing a model that allows exploration of hypotheses regarding the types of change in parameters such as natural mortality and stock-recruitment that might be expected. For example, in the case of the SNEMA yellowtail flounder, a BH model with a change in steepness or asymptotic level of recruitment following 1990 could be incorporated in the ASAP model, with the parameter describing the change in steepness or asymptotic recruitment being estimated when the model is fitted to the data. Such a model would provide greater ability to characterize the uncertainty associated with the possible change in the stock-recruitment relationship and allow determination of BRPs and stock status that take that uncertainty into account. By employing such a model, the need to consider two alternative recruitment scenarios, *i.e.,* low recruitment associated with either low SSB or reduced productivity, would be avoided.

A major concern of the current assessment for GoM/GB Atlantic herring is that, other than through the use of sensitivity runs, the uncertainty associated with the estimate of the post-1995 percentage increase in natural mortality is not carried through into the assessments of BRPs and determination of stock status. Further refinement of the time series of annual estimates of consumption of Atlantic herring, and particularly characterization of the uncertainty of the annual estimates, and continued development of an assessment model that incorporates this source of natural mortality are recommended, as this would facilitate assessment of the uncertainty associated with the consumption of Atlantic herring by predators.

Inadequate detail was provided in the assessment reports of the statistical designs of the bottom trawl surveys and the protocols used to obtain biological samples for these surveys and for the commercial catch data. Descriptions of the methods used to expand the collected data to estimates of catch at age or to combine the survey data from different samples to form a single index of abundance or overall age composition for each survey were either missing or inadequate.

The SARC meeting facilities and logistical support for the meeting were excellent. Information Technology support during the meeting ensured that access was available to the file server. The FTP site provided convenient access to electronic copies of background papers and assessment documents. The assistance provided by the rapporteurs, who recorded details of the discussions during the review, was greatly appreciated. It was pleasing to find that access to the meeting was available through WebEx, allowing a broader audience to participate.

6. References

Hoenig, J. M. 1983. Empirical use of longevity data to estimate mortality rates. *Fishery Bulletin*, **82**: 898-903.

McAllister, M. K., and Ianelli, J. N. 1997. Bayesian stock assessment using catch-age data and the sampling/importance resampling algorithm. *Canadian Journal of Fisheries and Aquatic Science*, **54**: 284–300.

Mertz, G., and Myers, R. A. 1996. Influence of fecundity on recruitment variability of marine fish. *Can. J. Fish. Aquat. Sci.*, **53**: 1618 - 1625.

Appendix 1: Bibliography of all material provided

Yellowtail Flounder

Background Papers

1 Wood, A. D., Cadrin, S. X., Alade, L. A., Martins, D., Moser, J., & Westwood, A. D. 2012. Mortality and movement of yellowtail flounder, *Limanda ferruginea*, tagged off New England.

2 McBride, R., Press, Y., & Wuenschel, M. 2012. Classifying female yellowtail flounder maturity: comparing at-sea, macroscopic maturity classifications with results from a gonad histology method.

3 McElroy, W. D., Towle, E. K., Press, Y. K., McBride, R. S., & Wuenschel, M. J.. 2012. Comparison of fecundity among stocks of female yellowtail flounder, *Limanda ferruginea*.

4 McElroy, W. D., Press, Y. K., & Wuenschel, M. J. 2012. Reproductive effort as a predictor of the natural mortality rate for southern New England yellowtail flounder: the Gunderson method.

5 Alade, L., & Cadrin, S, 2012. A review of yellowtail flounder stock structure off New England.

6 Alade, L., Hart, D., & Legault, C. 2012. Influence of spatial stratification on discard estimation for the Southern New England Mid-Atlantic yellowtail flounder.

7 Alade. L. 2012. Southern New England-Mid Atlantic yellowtail length-weight relationship.

8 Barkley, A., & Cadrin, S. 2012. Results of the application of reflex action mortality predictors onboard commercial fishing vessels.

9 Gervelis, B., Burchard, K., & Hoey, J. 2011. A comparison of discard estimates of yellowtail flounder using study fleet self-reported data and NEFOP observer data.

10 Alade, L. A. 2012. Southern New England Mid-Atlantic yellowtail flounder commercial landings.

11 Hare, J., Richardson, D., & Mountain, D. 2012. Indices of the Mid-Atlantic cold pool and relationship to Southern New England/ Mid-Atlantic yellowtail flounder recruitment.

12 Richardson, D. E., Walsh, H., & Hare, J. 2012. Southern New England yellowtail flounder larval data.

13 Anonymous. 2012. Revised Spatial Stratification for estimating discards for SNEMA yellowtail flounder.

Previous Assessments

14 GARM. 2002. Southern New England yellowtail flounder.

15 GARM. 2005. Southern New England-Mid Atlantic yellowtail flounder.

16 GARM. 2008. Southern New England/Mid Atlantic yellowtail flounder.

17 Anonymous. 2003. Yellowtail Flounder Stock Structure. 36th SAW Advisory.

18 Northeast Fisheries Science Center. 2003. Report of the 36th Northeast Regional Stock Assessment Workshop (36th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref. Doc. 03-06; 453 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026.

19 Anonymous. 2003. Southern New England - Mid Atlantic Yellowtail

Flounder. 36th SARC Consensus Summary.

Working Papers

1 Southern Demersal Working Group, Stock Assessment Workshop (SAW 54). 2012. Southern New England Mid-Atlantic (*Limanda ferruginea*) stock assessment updated through 2011. A report of the Southern Demersal Working Group for SARC 54, June 5th – 9th, 2012. Northeast Fisheries Science Center. Woods Hole, MA. SAW/SARC 54.

2 Anonymous. 2012. Southern New England Yellowtail Flounder Assessment Summary for 2012. SAW/SARC 54.

Atlantic Herring

Background Papers

1 Shepherd, G., Cieri, M., Power, M., & Overholtz, W. 2009. Transboundary Resources Assessment Committee Gulf of Maine/Georges Bank Atlantic Herring Stock Assessment Update. TRAC Reference Document 2009/04.

2 Transboundary Resources Assessment Committee. 2009. Gulf Of Maine-Georges Bank Herring Stock Complex. Status Report 2009/04.

3 Transboundary Resource Assessment Committee. 2006. Gulf of Maine-Georges Bank Herring Stock Complex. Status Report 2006/01.

4 Gavaris, S. 2003. Transboundary Resource Assessment Committee (TRAC): Report of Meeting held 10-14 February 2003. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2003/014.

5 O'Boyle, R., & Overholtz, W. 2006. Proceedings of the Transboundary Resources Assessment Committee (TRAC). Benchmark Review of Stock Assessment Models for Gulf of Maine and Georges Bank Herring 2 – 5 May 2006 Woods Hole, Massachusetts.

6 O'Brien, L., & T. Worcester, T. 2008. Proceedings of the Transboundary Resources Assessment Committee (TRAC): Gulf of Maine/Georges Bank Herring, Eastern Georges Bank Cod and Haddock, Georges Bank Yellowtail Flounder. Report of Meeting held 8 - 11 June 2009 St. Andrews Biological Station, St. Andrews, New Brunswick, Canada.

7 Jech, J. M., & Stroman, F. 2012. Aggregative patterns of pre-spawning Atlantic herring on Georges Bank from 1999-2010. *Aquat. Living Resour.*, **25**: 1–14.

8 Makris, N. C., Ratilal, P., Symonds, D. T., Jagannathan, S., Lee, S., & Nero, R. W. Fish population and behavior revealed by instantaneous continental shelf–scale imaging. *Science*, **311**: 660-663.

9 Makris, N. C., Ratilal, P., Jagannathan, S., Gong, Z., Andrews, M., Bertsatos, I., Godø, O. R., Nero, R. W., & Jech, J. M. 2009. Supporting online material for "Critical population density triggers rapid formation of vast oceanic fish shoals". *Science*, **323**: 1734.

10 Link. J. S., & Almeida, F. P. 2000. An Overview and History of the Food Web Dynamics Program of the Northeast Fisheries Science Center, Woods Hole, Massachusetts. NOAA Technical Memorandum NMFS-NE-159. 60 pp.

11 Kanwit, J. K., & Libby, D. A. 2009. Seasonal movements of Atlantic herring (*Clupea harengus*): results from a four year tagging study

conducted in the Gulf of Maine and Southern New England. *J. Northw. Atl. Fish. Sci.*, **40**: 29–39.

12 Guan, W., Cao, J., Chen, Y., & Matthew Cieri, M. 2012. A simulation study to evaluate impacts of spatial structure of Atlantic herring fishery on retrospective errors in stock assessment.

13 Reid, R. N., Cargnelli, L. M., Griesbach, S. J., Packer, D. B., Johnson, D. L., Zetlin, C. A., Morse, W. W. & Berrien, P. L. 1999. Essential Fish Habitat Source Document: Atlantic Herring, *Clupea harengus*, Life History and Habitat Characteristics. NOAA Technical Memorandum NMFS-NE-126. 48 pp.

14 Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology. 49: 627–647.

15 Miller, T. J., Richardson, D. E., & Hare, J. A. In prep. Maximum likelihood estimation of larval production indices from length frequency and growth information collected on ichthyoplankton surveys.

16 Cao, J., Guan, W., Chen, Y., & Cieri, M. In prep. Evaluating factors influencing retrospective errors in estimating stock biomass for Atlantic herring *Clupea harengus*.

Working Papers

1 Anonymous. 2012. Stock Assessment of Atlantic Herring- Gulf of Maine/Georges Bank. SAW/SARC 54. June 5-9 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

1 Appendices 3-6
Anonymous. 2012. Stock Assessment of Atlantic Herring- Gulf of Maine/Georges Bank. Appendices 3-6. SAW/SARC 54. June 5-9 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

2 Anonymous. 2012. Atlantic Herring Assessment Summary for 2012.

CIE Reviewer's Independent Report on

54th Northeast Regional Stock Assessment Workshop

(SAW 54)

Prepared by

Neil Klaer

Prepared for

Center for Independent Experts (CIE)

Stock Assessment Review Committee (SARC) Meeting
5 - 9 June 2012
Northeast Fisheries Science Center
Wood's Hole, Massachusetts

Contents

b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-7).

2.1.9 TOR-9 Using simulation/estimation methods, evaluate consequences of alternative harvest policies in light of uncertainties in model formulation, presence of retrospective patterns, and incomplete information on magnitude and variability in M.

2.1.10 TOR-10 Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).

b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

2.1.11 TOR-11 For any research recommendations listed in recent peer reviewed assessment and review panel reports, review, evaluate and report on the status of those research recommendations. Identify new research recommendations.

2.2 Review findings by term of reference: SNE/Mid-Atlantic Yellowtail Flounder

2.2.1 TOR-1 Estimate landings and discards by gear type and where possible by fleet, from all sources. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.

2.2.2 TOR-2 Present the survey data being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, age-length data, etc.). Investigate the utility of commercial or recreational LPUE as a measure of relative abundance, and characterize the uncertainty and any bias in these sources of data.

2.2.3 TOR-3 Evaluate the validity of the current stock definition, and determine whether it should be changed. Take into account what is known about migration among stock areas.

2.2.4 TOR-4 Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-5), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

2.2.5 TOR-5 Investigate causes of annual recruitment variability, particularly the effect of temperature. If possible, integrate the results into the stock assessment (TOR-4).

2.2.6 TOR-6 State the existing stock status definitions for "overfished" and "overfishing". Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

2.2.7 TOR-7 Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model, should one be developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).

a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.

b. Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-6).

2.2.8 TOR-8 Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

a. Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment, and recruitment as a function of stock size).

b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.

c. Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.

2.2.9 TOR-9 Review, evaluate and report on the status of research recommendations listed in most recent peer reviewed assessment and review panel reports. Identify new research recommendations.

3 **Critique of the review process**
4 **References**

Appendix 1: Bibliography of materials provided for review
Appendix 2: A copy of the CIE Statement of Work
Appendix 3: List of participants

Executive Summary

The Stock Assessment Review Committee (SARC) of the 54th Stock Assessment Workshop (SAW 54) met at the Northeast Fisheries Science Center (NEFSC), Woods Hole, MA during 5th – 9th June 2012 to review Northeast regional benchmark stock assessments of Atlantic Herring *(Clupea harengus)* and Southern New England / Mid-Atlantic Bight Yellowtail Flounder (*Limanda ferruginea*). The SARC review panel (herein called the "Panel") consisted of Chair Robert O'Boyle (Beta Scientific Consulting, Canada) and three scientists representing the Center for Independent Experts (CIE): Chris Francis, Norm Hall and Neil Klaer.

The meeting format included presentations mixed with questions and open discussion. The Panel participated in the review of each term of reference. The meeting was open to the public and public comments were also accepted.

Findings by term of reference

2.1 Atlantic Herring

2.1.1 TOR-1 Successfully completed.

- The WG has made a good compilation of landings and discard data by fixed and mobile gear types from all sources, that extensive length and age sampling data of apparent good quality were available.

- It would be possible to develop alternative catch series that take some account of the uncertainty in the stock boundary.

2.1.2 TOR-2 Successfully completed.

- Reasonable justification was given to use the NMFS spring, fall, and shrimp bottom trawl surveys, and not the winter, larval, and state-run surveys as abundance indices for the stock assessment.

- Commercial LPUE was discounted as a usable index of abundance because of the effect of fishing regulations on locations fished, hyperstability, and the difficulty of identifying "herring trips" because of target switching within trips.

2.1.3 TOR-3 Successfully completed.

- The NEFSC fall acoustic survey was not used as it was seen to cover a limited spatial area that was not representative of the entire stock.

2.1.4 TOR-4 Successfully completed.

- Given the requirement for an assessment of the stock, one major sub-stock confined within the Gulf of Maine/Georges Bank region and uncertainty in the level of mixing of the other two sub-stocks, the Panel agreed with the WG decision to maintain the current stock definition for management purposes.

2.1.5 TOR-5 Successfully completed.

- The base case assessment model with a 50% increase in *M* since 1996 was accepted as the most plausible model for management purposes.

- Key reasons for acceptance of this model with an *M* increase were independent information about a likely increase in Atlantic Herring consumption by predators since about 1996, and the resolution of a retrospective pattern that has caused major concern for the previous Atlantic Herring assessment.

- More work could be done to rule out other possible causes of the retrospective pattern.

2.1.6 TOR-6 Successfully completed.

- The *M* change was justified because of the wealth of available annual information from stomach contents for a wide range of herring predators (fish, mammals, other), and annual abundance estimates of those predators.

- Uncertainty in the level of increase in *M* was examined but others such as the start year for the *M* change, the shape of the *M* adjustment other than a step function, and development of alternative plausible *M* scenarios based on uncertainty in consumption were not.

2.1.7 TOR-7 Successfully completed.

- The Panel concluded that the BRPs calculated using the base model were appropriate for the immediate future (3 to 5 years).

2.1.8 TOR-8 Successfully completed.

- Status was determined using the new model and updated data under a range of *M* (both age- and time–specific and post-1995) and BH steepness options and, except for the unlikely case where steepness was assumed to be 0.35, the current status was estimated to be not overfished and overfishing was not occurring.

2.1.9 TOR-9 Not completed, but some initial work underway.

- Alternative operating model scenarios could be developed that make various assumptions about the mixing and stock boundary effects of sub-stocks and the magnitude and variability of an *M* change.

- Specification of the management objectives and performance measures requires considerable input from management and stakeholders.

2.1.10 TOR-10 Successfully completed.

- Projection methods were sound, and were applied to a wide range of scenarios that successfully spanned the plausible range of uncertainty. Key sources of this stock's vulnerability to becoming overfished were identified and well described.

2.1.11 TOR-11 Successfully completed.

- The Panel commented on the priority of research items listed by the WG, and also made additional research recommendations.

2.2 Yellowtail Flounder

2.2.1 TOR-1 Successfully completed.

- The procedure used by the WG to produce best annual estimates of total landings and discards for the stock was well justified.

- Alternative plausible catch histories should be developed so that this source of uncertainty can be carried into future stock assessments.

- A summary table of available age, length and weight samples by year should be prepared as part of the assessment documentation.

2.2.2 TOR-2 Successfully completed.

- The Panel endorsed the use of the NMFS spring, fall, and winter surveys, and the larval survey, and, because of poor sampling, the exclusion of the southern strata when calculating abundance indices for the winter survey.

- Uncertainty in survey calibration factors was not carried through into the stock assessment.

- Commercial LPUE is unlikely to provide a useful index of abundance due to changes in management regulations, changes in reporting methodology, and the change of the fishery from directed to mostly bycatch.

2.2.3 TOR-3 Successfully completed.

- Available evidence makes a strong case for the southern New England (mid-Atlantic Bight) SNEMA region being a single stock for management purposes.

2.2.4 TOR-4 Successfully completed.

- The base case is an adequate basis for management decisions.

- The statistical catch at age ASAP model is appropriate given the data available.

- There were some concerns that the data weightings were somewhat ad-hoc, and the Panel provided some recommendations on weighting procedures.

- Major uncertainties in the assessment were well characterised by the MCMC analyses and alternate model runs. However there were some uncertainties that were not explored as part of the assessment including survey calibrations, catch history (particularly discards), and base natural mortality rate.

2.2.5 TOR-5 Successfully completed.

- One hypothesis for the recruitment pattern shown by the stock assessment is that the low recruitment levels since 1990 were influenced by a shift in environmental conditions.

- A number of sources of information were shown at the SARC that documented long-term trends or highly variable oceanographic conditions that could influence

Yellowtail Flounder productivity. None, however, showed a pattern that indicated a major shift since about 1990.

2.2.6 TOR-6 Successfully completed.

- In calculating BRPs, the WG considered two alternative scenarios: ‘two-stanza’ which links the drop in recent recruitment to a decrease in SSB and a ‘recent’ scenario in which the drop was due to a productivity shift caused by unknown environmental changes.

- The Panel used a weight of evidence approach to conclude that the evidence was 60:40 in favor of the ‘recent’ scenario, although both scenarios were included in advice to management.

- Values of MSY and B_{MSY} calculated under the two-stanza scenario were surprisingly different from those calculated (during the review meeting) using a modelling approach (with a BH stock-recruit relationship) that is more conventional in other fora. This difference needs further investigation.

2.2.7 TOR-7 Successfully completed.

- When evaluated against the BRPs derived from the two-stanza recruitment scenario, the stock is found to be overfished but when evaluated using the BRPs derived from the recent recruitment scenario, the stock is not overfished and the stock is rebuilt.

- While the Panel considers that the drop in recent recruitment is more likely than not due to a productivity change (see TOR-6), the alternate hypothesis cannot be ruled out.

- Additional scientific advice that could be provided to management in the current situation of dual plausible stock situations is an analysis of the risk to the stock or catches of making an incorrect decision.

2.2.8 TOR-8 Successfully completed.

- The projection methods used in the assessment were sound, and were applied to two alternative scenarios: ‘recent’ and ‘two-stanza’.

2.2.9 TOR-9 Successfully completed.

- The Panel commented on the priority of research items listed by the WG, and also made additional research recommendations.

- I also suggest that an analysis of the consequences of acceptance of the wrong scenario can be undertaken as part of future research.

1.1 Background

The Stock Assessment Review Committee (SARC) of the 54th Stock Assessment Workshop (SAW 54) met at the Northeast Fisheries Science Center (NEFSC), Woods Hole, MA during 5th – 9th June 2012 to review Northeast regional benchmark stock assessments of Atlantic Herring *(Clupea harengus)* and Southern New England / Mid-Atlantic Bight Yellowtail Flounder (*Limanda ferruginea*), guided by the SAW 54 Terms of Reference (Annex 2 of the SAW 54 Statement of Work provided below).

The SARC review panel consisted of Chair Robert O'Boyle (Beta Scientific Consulting, Canada) and three scientists representing the Center for Independent Experts (CIE): Chris Francis, Norm Hall and Neil Klaer.

The SARC was assisted by the NEFSC SAW Chairman, James Weinberg, Anne O'Brien, and Paul Rago (NEFSC). Documentation for the herring assessment was prepared by the NEFSC Herring Working Group (HWG), and the presentations at the meeting were made by Jon Deroba (NEFSC). Documentation for the yellowtail assessment was prepared by the NEFSC Southern Demersal Working Group (SDWG), and the presentation at the meeting was made by Larry Alade (NEFSC). The rapporteurs who recorded the discussion to assist the Panel in its deliberations were Toni Chute (Atlantic Herring) and Jessica Blaylock (Yellowtail Flounder).

1.2 Review of Activities

The SARC met at Woods Hole from Tuesday 5th to Saturday 9th June 2012, the agenda of which is summarized in Table 1 (see full agenda in annex 3 of the SAW Statement of Work below).

Table 1. Summary of SARC/SAW 54 Agenda during 5th – 9th June 2012

	5th June	6th June	7th June	8th June	9th June
	Tuesday	**Wednesday**	**Thursday**	**Friday**	**Saturday**
Morning	Panel meeting (10:30)	Herring assessment	Assessment revisits	Summary Report Review	Panel Report Discussion (closed meeting)
Afternoon	Herring Assessment	SNE Yellowtail assesment	Assessment revisits	Summary Report Review	
			Summary Report Review	Panel Report Discussion (closed meeting)	

Each reviewer on the Panel was assigned to individual Terms of Reference (TORs) for each species to compile summary points and to help in the preparation of the summary report. I was assigned to TOR-1 landings and effort, TOR-4 stock definition and TOR-11 research recommendations for Atlantic Herring, and TOR-1 landings and effort, TOR-3 stock definition and TOR-9 research recommendations for Yellowtail Flounder.

The Panel devoted Friday afternoon and Saturday morning to distilling and combining summary points for each stock's Terms of Reference as well as observations on the SARC process. It was agreed that each panelist would use these points to draft a section of the summary report, which was then to be compiled and edited by the SARC Chair. There were no disagreements among the reviewers on the contents of the summary report, so my own report here reflects the contents of that report, and provides some additional information.

ATLANTIC HERRING

2.1 Findings by term of reference

2.1.1 TOR-1 Estimate catch from all sources including landings and discards. Describe the spatial distribution of fishing effort. Characterize uncertainty in these sources of data.

The Panel concluded that this term of reference had been successfully completed and I agree with the points made in the summary report.

The Working Group (WG) has made a good compilation of landings and discard data by fixed and mobile gear types from all sources, extensive length and age sampling data of apparent good quality were available, and observer coverage of the commercial fishery has been sufficient particularly recently to give confidence in discard estimates.

Some account of uncertainty in catch and discarding was made in the stock assessment through a coefficient of variation (CV) on the catch, and tuning of effective sample sizes on proportions of catch at age. However, greater uncertainty of catches from the stock may be due to the decision on placement of the stock boundary. In particular, the influence of catches in the Scotian Shelf region on the stock is unknown. It would be possible to develop alternative catch series that take some account of the uncertainty in the stock boundary.

Alternative catch series may be developed that attempt to bracket the best estimate, so providing possible low and high alternatives. In the case here, perhaps the most important alternative series would be a high catch one that includes some plausible proportion of catches from the Scotian Shelf.

2.1.2 TOR-2 Present the survey data being used in the assessment (e.g., regional indices of abundance, recruitment, state surveys, larval surveys, age-length data, predator consumption rates, etc.). Investigate the utility of commercial LPUE as a measure of relative abundance, and characterize the uncertainty and any bias in these sources of data.

The Panel concluded that this term of reference had been successfully completed and I agree with the points made in the summary report.

The WG provided reasonable justification to use the NMFS spring, fall, and shrimp bottom trawl surveys, and not the winter, larval, and state-run surveys as abundance indices for the stock assessment. The use of commercial LPUE as an index of abundance was discounted because of the effect of fishing regulations on locations fished, hyperstability, and the difficulty of identifying “herring trips” because of target switching within trips. The Panel agreed that LPUE would not provide a useful index of abundance.

It is a general disappointment when long-term survey series are broken due to fishing gear changes. For Atlantic Herring those included changes to the nets, trawl door and vessels within key series. Calculation of calibration factors to allow some series continuity is a particular issue and problem for the assessment of Atlantic Herring.

2.1.3 TOR-3 Evaluate the utility of the NEFSC fall acoustic survey to the stock assessment of herring. Consider degree of spatial and temporal overlap between the survey and the stock. Compare acoustic survey results with measures derived from bottom trawl surveys.

The Panel concluded that this term of reference had been successfully completed and I agree with the points made in the summary report.

The Panel concurred with the decision not to use the NEFSC fall acoustic survey as it was seen to cover a limited spatial area that was not representative of the entire stock.

2.1.4 TOR-4 Evaluate the validity of the current stock definition, and determine whether it should be changed. Take into account what is known about migration among stock areas.

The Panel concluded that this term of reference had been successfully completed and I agree with the points made in the summary report.

There are at least three major sub-stocks of herring within the defined boundaries of the management region encompassing the Gulf of Maine/Georges Bank stock complex, with known mixing at the northern and southern boundaries. Given the requirement for an assessment of the stock, one major sub-stock confined within the Gulf of Maine/Georges Bank region and uncertainty in the level of mixing of the other two sub-stocks, the Panel agrees with the WG decision to maintain the current stock definition for management purposes.

Although the WG has stated that separation of catches and catch composition information by sub-stocks is not possible at present, the Panel believes that such data separation is one of the major barriers to improving the stock assessment, and agrees with the WG that future research should be directed towards data separation by sub-stock. For example, movement rates from tagging studies may be used to create generalized sub-stock mixing rates which could then be used in a multi-substock assessment, or used directly to separate assessment input data by sub-stock.

Sub-stock mixing at the southern and particularly northern stock boundaries introduces one of the major uncertainties into the stock assessment. Scenarios should be developed that account for this uncertainty that can be carried through to the stock assessment. One such approach is to develop alternative catch scenarios as discussed under TOR-1.

2.1.5 TOR-5 Estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series (integrating results from TOR-6), and estimate their uncertainty. Include a historical retrospective analysis to allow a comparison with previous assessment results and previous projections.

The Panel concluded that this term of reference had been successfully completed and I agree with the points made in the summary report.

The base case assessment model with a 50% increase in *M* since 1996 was accepted as the most plausible model for management purposes.

Key reasons for acceptance of this model with an *M* increase were independent information about a likely increase in Atlantic Herring consumption by predators during that time, and the resolution of a retrospective pattern that caused major concern for the previous Atlantic Herring assessment. The case made for the increased *M* due to increased consumption was convincing (see TOR-6) and led the Panel to accept that scenario as likely.

However, although the increased *M* resolved the retrospective pattern, I believe that more work could be done on this aspect to provide additional confidence in the increased *M* scenario.

The existing retrospective pattern was characterized by an apparent overestimation of recent abundance. To resolve the pattern, a mechanism to reduce the number of fish in the recent population was required. A change in *M* was selected as a likely mechanism due to evidence from consumption information. Other causes or contributors to the retrospective pattern are possible including underestimated fishing mortality or change in survey *q* values (e.g. see Mohn 1999). These other possible causes should be investigated and ruled out if possible. A simple investigation could be made using the assessment model alone to determining what level of change in these other elements would resolve the retrospective pattern, and then to provide a judgment of whether such a scenario is plausible. An improved evaluation would be through development of alternative plausible scenarios as modifications to the operating model in a management strategy evaluation (MSE) (see TOR-9).

A common means for examining major uncertainties in stock assessments that has become routine in recent years is sensitivity analysis – a systematic examination of changes to all major assumptions from the base case (e.g. base *M* and *h* values, and relative likelihood weighting given to different abundance indices and age/length composition). For Atlantic herring, an additional specific assumption requiring examination was the level of *M* increase. These individual changes are normally carried through to the management advice that follows (e.g. resulting F_{msy} and projected catch values). I would also add a table of likelihood components for each sensitivity run so that changes to the model fit can also be examined. Many of these sensitivities were examined during the

SARC Review, allowing the Panel to agree that these uncertainties had been examined under this TOR, but those examinations might best be carried out by the WG as a routine component of the stock assessment.

2.1.6 TOR-6 Consider the implications of consumption of herring, at various life stages, for use in estimating herring natural mortality rate (*M*) and to inform the herring stock-recruitment relationship. Characterize the uncertainty of the consumption estimates. If possible integrate the results into the stock assessment.

The Panel concluded that this term of reference had been successfully completed and I agree with the points made in the summary report.

Estimates of consumption of Atlantic herring were used to inform the decision to increase post-1995 natural mortality in the base assessment model. It is unusual to consider annual consumption as a mechanism for adjustment of natural mortality in a stock assessment, but it was justified in this case because of the wealth of available annual information from stomach contents for a wide range of herring predators (fish, mammals, other), and annual abundance estimates of those predators. Most of the signal in the annual consumption of herring derived from stomach contents rather than predator abundance.

Despite high uncertainty in the resulting estimates of annual herring consumption, the data provided good evidence that consumption of herring had increased since 1996. The Panel noted, however, that (1) the later peak in the time series of annual consumption estimates was driven by the very high abundances of two individual predator species, but similar high levels of abundance for those species were not present in adjacent years; (2) abundance estimates of some predator species were calculated from swept area calculations, rather than assessment models; (3) the consumption estimates used in this assessment were likely to be underestimates of total consumption; and (4) estimates of consumption were included in an exploratory run of the assessment model but did not improve model fit.

Uncertainty in the level of increase in *M* was examined during the SARC meeting, primarily to determine whether there was good justification for using 50%. Other uncertainties, such as the start year for the *M* change, the shape of the *M* adjustment other than a step function, and development of alternative plausible *M* scenarios based on uncertainty in consumption were not examined and are noted for future research.

2.1.7 TOR-7 State the existing stock status definitions for “overfished” and “overfishing”. Then update or redefine biological reference points (BRPs; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on

the scientific adequacy of existing BRPs and the "new" (i.e., updated, redefined, or alternative) BRPs.

The Panel concluded that this term of reference had been successfully completed and I agree with the points made in the summary report.

The Panel concluded that the BRPs calculated using the base model, i.e., with a base level of *M* = 0.3 and a post-1995 increase in *M* of 50% due to consumption by predators, was appropriate for the immediate future (3 to 5 years). The Panel recommended, however, that monitoring of predation be continued due to large uncertainty in the assumption that consumption of Atlantic herring by predators would remain at current levels in the longer-term.

2.1.8 TOR-8 Evaluate stock status with respect to the existing model (from previous peer reviewed accepted assessment) and with respect to a new model, should one be developed for this peer review. In both cases, evaluate whether the stock is rebuilt (if in a rebuilding plan).

- **When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.**
- **Then use the newly proposed model and evaluate stock status with respect to "new" BRPs and their estimates (from TOR-7).**

The Panel concluded that this term of reference was successfully completed and I agree with the points made in the summary report.

Status was determined using the new model and updated data under a range of *M* (both age and time – specific and post-1995) and BH steepness options and, except for the unlikely case where steepness was assumed to be 0.35, the current status was estimated to be not overfished and overfishing was not occurring.

2.1.9 TOR-9 Using simulation/estimation methods, evaluate consequences of alternative harvest policies in light of uncertainties in model formulation, presence of retrospective patterns, and incomplete information on magnitude and variability in *M*

The Panel concluded that this term of reference was not completed, but that some initial work was underway. I agree with the points made in the summary report.

Alternative operating model scenarios could be developed for Atlantic herring that make various assumptions about the mixing and stock boundary effects of sub-stocks. Those could encompass the effects of fishing outside the management boundary, and also uncertainty and representativeness of data collections within the boundary. The magnitude and variability of an *M* change is another major component to explore.

While a range of alternative plausible "states of nature" can be hypothesised and implemented as operating model scenarios, I agree with the Panel that the specification of the management objectives and performance measures requires considerable input from management and stakeholders.

2.1.10 TOR-10 Develop approaches and apply them to conduct stock projections and to compute the pdf (probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).

- **Provide numerical annual projections (3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).**
- **Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.**
- **Describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming overfished, and how this could affect the choice of ABC.**

The Panel concluded that this term of reference was successfully completed.

The projection methods were sound, and were applied to a wide range of scenarios that successfully spanned the plausible range of uncertainty. Key sources of this stock's vulnerability to becoming overfished (e.g., contributions from other herring stocks, uncertainty about the strength of the 2008 year class and the persistence of high natural mortality) were identified and well described.

2.1.11 TOR-11 For any research recommendations listed in recent peer reviewed assessment and review panel reports, review, evaluate and report on the status of those research recommendations. Identify new research recommendation

The Panel concluded that this term of reference had been successfully completed. I agree on the comments made in the summary report about the priority of research items listed by the WG, and also agree with the additional research recommendations made by the Panel.

I have added some detail here to Panel recommendation (t):

> t. Using simulation/estimation methods, evaluate consequences of alternative harvest policies in light of uncertainties in model formulation, presence of retrospective patterns, and incomplete information on the magnitude and variability in *M* (TOR-9). Uncertainties to be examined for *M* could include the start year for the *M* change, the shape of the *M* adjustment other than a step function, and alternative plausible *M* scenarios based on consumption data.

3 Critique of the review process

The Panel reached summary on all Terms of Reference for each stock. It acknowledges the significant work that the two assessment working groups had undertaken to prepare for the SARC review. It also appreciates the professionalism and cooperation of NEFSC staff at the SARC meeting, which significantly assisted the peer review. Notwithstanding this, during the course of the review, issues came to the attention of the Panel that were not specific to the assessment of either species, resolution of which would assist future SARC reviews. These relate to both the terms of reference of reviews and the presentation of assessment results.

The terms of reference of the herring and yellowtail assessments required a review of their stock definitions. These were conducted during the data meetings of each species, the results of which were brought forward to the assessment meeting. Changes in stock definition have consequences throughout the management system and should not be undertaken without significant consideration of all sources of information. One could expect that there would be considerable reluctance to change stock definition without substantial evidence to the contrary. The Panel considers that reviews of stock definition would more productively be undertaken outside of the normal assessment process and on a schedule that would allow significant changes if these were felt warranted. The review of stock definition needs to highlight the uncertainties in the interactions amongst populations that might influence the interpretation of data during an assessment. This review also needs to determine the catch and indices appropriate for the stock(s) in question.

The term of reference for each stock included review of the data (e.g. catch, indices) to be used in the assessments. Consideration should be given to formally separating the data review from that of the assessment, similar to the SEDAR process conducted by the NMFS Southeast Fisheries Science Center. This would allow the peer review of the assessment to fully devote its attention to determination of stock status, reference points and projections. The data reviews could be undertaken on groups of species (e.g. groundfish, pelagics) to obtain data perspectives across stocks. These reviews would indicate the relative reliability of various datasets (e.g., survey indices) and clearly specify which data and their uncertainties should be brought forward to the assessment review.

While it was evident that the HWG and SDWG had spent considerable time preparing the documents and presentations for the SARC 54, there was unevenness in the relative content of each. In one stock, information was summarized in the working papers that were not summarized in the presentations at the SARC review, while in the other stock, the reverse was the case. Greater detail was sometimes available in the presentation than was in the working papers. It would assist future peer reviews if the evidence supporting the conclusions of a working group were both presented in the assessment report and the presentation such that panel reviewers may assess whether or not the conclusions are justified.

On some of the figures describing current stock status, both the posterior distribution of the relevant indicator (e.g., spawning stock biomass), as determined through an MCMC process, and point estimates were displayed. Where MCMC was used to provide distributions of current stock status, these should be used to provide metrics and their uncertainty required by the management system. In this case, coefficients of variations based on the Hessian matrix are not required to be reported.

In both assessments, historical trends in spawning stock biomass and fishing mortality were illustrated along with current estimates of the biological reference points (BRPs). However, as highlighted by both assessments, temporal changes in population processes can dramatically change the BRPs. Changes in growth and fishery selectivity can also change the BRPs. There are a number of ways temporal changes in BRPs can be displayed (annually, smoothed over a number of years, by decade, etc.). It would be useful for the NEFSC to develop a policy for the estimation and presentation of temporal changes in BRPs to both avoid future confusion and promote transparent communication with stakeholders and managers on long-term productivity and fisheries changes.

4 References

Krimsky, S. 2005. The Weight of Scientific Evidence in Policy and Law. American Journal of Public Health, Suppl. 1, 95:S1.

Mohn, R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES Journal of Marine Science, 56: 473–488.

Wayte, S.E. in review. Management implications of including a climate-induced recruitment shift in the stock assessment for jackass morwong (*Nemadactylus macropterus*) in south-eastern Australia. Fisheries Research, special issue on Stock Synthesis.

Appendix 1. Bibliography

Yellowtail Flounder

Background Papers

1 Wood, A. D., Cadrin, S. X., Alade, L. A., Martins, D., Moser, J., & Westwood, A. D. 2012. Mortality and movement of yellowtail flounder, *Limanda ferruginea*, tagged off New England.

2 McBride, R., Press, Y., & Wuenschel, M. 2012. Classifying female yellowtail flounder maturity: comparing at-sea, macroscopic maturity classifications with results from a gonad histology method.

3 McElroy, W. D., Towle, E. K., Press, Y. K., McBride, R. S., & Wuenschel, M. J. 2012. Comparison of fecundity among stocks of female yellowtail flounder, *Limanda ferruginea*.

4 McElroy, W. D., Press, Y. K., & Wuenschel, M. J. 2012. Reproductive effort as a predictor of the natural mortality rate for southern New England yellowtail flounder: the Gunderson method.

5 Alade, L., & Cadrin, S, 2012. A review of yellowtail flounder stock structure off New England.

6 Alade, L., Hart, D., & Legault, C. 2012. Influence of spatial stratification on discard estimation for the Southern New England Mid-Atlantic yellowtail flounder.

7 Alade, L. 2012. Southern New England-Mid Atlantic yellowtail length-weight relationship.

8 Barkley, A., & Cadrin, S. 2012. Results of the application of reflex action mortality predictors onboard commercial fishing vessels.

9 Gervelis, B., Burchard, K., & Hoey, J. 2011. A comparison of discard estimates of yellowtail flounder using study fleet self-reported data and NEFOP observer data.

10 Alade, L. A. 2012. Southern New England Mid-Atlantic yellowtail flounder commercial landings.

11 Hare, J., Richardson, D., & Mountain, D. 2012. Indices of the Mid-Atlantic cold pool and relationship to Southern New England/ Mid-Atlantic yellowtail flounder recruitment.

12 Richardson, D. E., Walsh, H., & Hare, J. 2012. Southern New England yellowtail flounder larval data.

13 Anonymous. 2012. Revised Spatial Stratification for estimating discards for SNEMA yellowtail flounder.

Previous Assessments

14 GARM. 2002. Southern New England yellowtail flounder.

15 GARM. 2005. Southern New England-Mid Atlantic yellowtail flounder.

16 GARM. 2008. Southern New England/Mid Atlantic yellowtail flounder.

17 Anonymous. 2003. Yellowtail Flounder Stock Structure. 36th SAW Advisory.

18 Northeast Fisheries Science Center. 2003. Report of the 36th Northeast Regional Stock Assessment Workshop (36th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref. Doc. 03-06; 453 p. Available from: National Marine Fisheries

19 Anonymous. 2003. Southern New England - Mid Atlantic Yellowtail Flounder. 36th SARC Consensus Summary.

Working Papers

1 Southern Demersal Working Group, Stock Assessment Workshop (SAW 54). 2012. Southern New England Mid-Atlantic (*Limanda ferruginea*) stock assessment updated through 2011. A report of the Southern Demersal Working Group for SARC 54. June 5th – 9th, 2012. Northeast Fisheries Science Center. Woods Hole, MA. SAW/SARC 54.

2 Anonymous. 2012. Southern New England Yellowtail Flounder Assessment Summary for 2012. SAW/SARC 54.

Atlantic Herring

Background Papers

1 Shepherd, G., Cieri, M., Power, M., & Overholtz, W. 2009. Transboundary Resources Assessment Committee Gulf of Maine/Georges Bank Atlantic Herring Stock Assessment Update. TRAC Reference Document 2009/04.

2 Transboundary Resources Assessment Committee. 2009. Gulf Of Maine-Georges Bank Herring Stock Complex. Status Report 2009/04.

3 Transboundary Resource Assessment Committee. 2006. Gulf of Maine-Georges Bank Herring Stock Complex. Status Report 2006/01.

4 Gavaris, S. 2003. Transboundary Resource Assessment Committee (TRAC): Report of Meeting held 10-14 February 2003. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2003/014.

5 O'Boyle, R., & Overholtz, W. 2006. Proceedings of the Transboundary Resources Assessment Committee (TRAC). Benchmark Review of Stock Assessment Models for Gulf of Maine and Georges Bank Herring 2 – 5 May 2006 Woods Hole, Massachusetts.

6 O'Brien, L., & T. Worcester, T. 2008. Proceedings of the Transboundary Resources Assessment Committee (TRAC): Gulf of Maine/Georges Bank Herring, Eastern Georges Bank Cod and Haddock, Georges Bank Yellowtail Flounder. Report of Meeting held 8 - 11 June 2009 St. Andrews Biological Station, St. Andrews, New Brunswick, Canada.

7 Jech, J. M., & Stroman, F. 2012. Aggregative patterns of pre-spawning Atlantic herring on Georges Bank from 1999-2010. *Aquat. Living Resour.*, **25**: 1–14.

8 Makris, N. C., Ratilal, P., Symonds, D. T., Jagannathan, S., Lee, S., & Nero, R. W. Fish population and behavior revealed by instantaneous continental shelf–scale imaging. *Science*, **311**: 660-663.

9 Makris, N. C., Ratilal, P., Jagannathan, S., Gong, Z., Andrews, M., Bertsatos, I., Godø, O. R., Nero, R. W., & Jech, J. M. 2009. Supporting online

Bertsatos, I., Godø, O. R., Nero, R. W., & Jech, J. M. 2009. Supporting online material for "Critical population density triggers rapid formation of vast oceanic fish shoals". *Science*, **323**: 1734.

10 Link. J. S., & Almeida, F. P. 2000. An Overview and History of the Food Web Dynamics Program of the Northeast Fisheries Science Center, Woods Hole, Massachusetts. NOAA Technical Memorandum NMFS-NE-159. 60 pp.

11 Kanwit, J. K., & Libby, D. A. 2009. Seasonal movements of Atlantic herring (*Clupea harengus*): results from a four year tagging study conducted in the Gulf of Maine and Southern New England. *J. Northw. Atl. Fish. Sci.*, **40**: 29–39.

12 Guan, W., Cao, J., Chen, Y., & Matthew Cieri, M. 2012. A simulation study to evaluate impacts of spatial structure of Atlantic herring fishery on retrospective errors in stock assessment.

13 Reid, R. N., Cargnelli, L. M., Griesbach, S. J., Packer, D. B., Johnson, D. L., Zetlin, C. A., Morse, W. W. & Berrien, P. L. 1999. Essential Fish Habitat Source Document: Atlantic Herring, *Clupea harengus*, Life History and Habitat Characteristics. NOAA Technical Memorandum NMFS-NE-126. 48 pp.

14 Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology. 49: 627–647.

15 Miller, T. J., Richardson, D. E., & Hare, J. A. In prep. Maximum likelihood estimation of larval production indices from length frequency and growth information collected on ichthyoplankton surveys.

16 Cao, J., Guan, W., Chen, Y., & Cieri, M. In prep. Evaluating factors influencing retrospective errors in estimating stock biomass for Atlantic herring *Clupea harengus*.

Working Papers

1 Anonymous. 2012. Stock Assessment of Atlantic Herring- Gulf of Maine/Georges Bank. SAW/SARC 54. June 5-9 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

1 Appendices 3-6

Anonymous. 2012. Stock Assessment of Atlantic Herring- Gulf of Maine/Georges Bank. Appendices 3-6. SAW/SARC 54. June 5-9 2012, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

2 Anonymous. 2012. Atlantic Herring Assessment Summary for 2012.