B. Stock Assessment of Georges Bank (GBK) Winter Flounder for 2011

Executive Summary

Term of Reference 1: ***Estimate catch from all sources, including landings and discards. Characterize the uncertainty in these sources of data.***

Catches were dominated by landings from the U.S. groundfish bottom trawl fleet during 1964-2010. Since 1964, total landings have been predominately from the U.S groundfish trawl fishery, but landings have also been reported for the Canadian groundfish trawl fisheries, as bycatch in the haddock and cod fisheries. Total landings, mainly from the U.S. and USSR fleets, increased during 1965-1972 to a time series peak of 4,509 mt. During 1970-1993, Canadian landings generally comprised a low percentage (1-2 %) of the total landings, but thereafter increased from 6% in 1994 to a peak of 24 % in 2001 then declined to low levels since then.

Total landings increased during 1964-1972, reaching a peak of 4,509 mt in 1972, then declined to1,892 mt in 1976. A sustained period of high landings occurred during 1977-1984, ranging from 3,061-4,009 mt. After 1984, landings gradually declined to the lowest level in the time series, 783 mt in 1995, but then increased again to 3,139 mt in 2003. Thereafter, landings declined rapidly and reached the second lowest level on record in 2007 (807 mt). During the time period included in the stock assessment model, 1982-2010, total landings averaged 1,950 mt and were slightly below this average in 2009 (1,670 mt) and 2010 (1,297 mt).

During the assessment period, 1982-2010, discards of winter flounder on Georges Bank were higher in the U.S. fisheries prior to 1991 (i.e., primarily from the large mesh (≥ 5.5 in. codend mesh size) fleet during 1964-1975 and the scallop dredge fleet during 1976-2010) and were higher in the Canadian scallop dredge fishery thereafter. Discards of winter flounder by the Canadian groundfish trawl fleet were not available. Total discards were much higher during 1982-1991, than thereafter, but total discards slowly increased between 1995 and 2010. Total discards averaged 15% of the total landings during 1982-2010.

Catches increased during 1964-1972, reaching a peak of 4,608 mt in 1972, but then declined to 2,034 mt in 1976. Catches subsequently increased to 4,290 mt in 1981 then gradually declined to a time series low of 842 mt in 1995. Catches increased to 3,328 mt in 2003 then declined to 1,039 mt in 2007, followed by a slight increase to 2,013 mt in 2009. Total catch in 2010 was 1,544 mt.

Similar to the most recent assessment, in 2008, the stock was assessed using an ADAPT VPA model. Components of the catch-at-age (CAA) consisted of the combined U.S. and Canadian landings-at-age and discards-at-age for the U.S. large-mesh and small-mesh bottom trawl fleets plus the U.S. and Canadian scallop dredge fleets, during 1982-2010, for ages 1-7+. During 1982-1984, the CAA contained a broad range of ages, but was dominated by ages 2-5 and had the highest numbers of fish aged 6 and older. The CAA changed from this more stable age composition to one dominated by ages 2-4, during 1985-1996. During 2000-2005, the catch composition changed back to a predominance of age 3-5 fish and contained more older fish (ages 6 and older), but at the higher levels observed during 1980-1984. The catches were dominated by age 2-4 fish during 2008-2010 as the 2006 year class was harvested by the fishery.
Mean weights-at-age in the catch remained relatively stable during 1985-1996 across most ages, but then declined to a lower level during 1997-2001, for ages 3-5 possibly due, in part, to poor sampling of large fish during part of this time period (Figure B17, Table B16). Mean weights-at-age reached their highest

levels during 2003-2007, but then declined through 2010 to some of their lowest levels since 1982.

Term of Reference 2: ***Present survey data being considered and/or used in the assessment (e.g., regional indices of abundance, recruitment, state and other surveys, age-length data, etc.). Characterize uncertainty in these sources of data.***

Minimum population sizes estimated from the Canadian and U.S. spring surveys and the U.S. fall surveys, for ages 1-7+ during 1982-2010,were included in the VPA model. A fourth order polynomial model was fit to the U.S. survey data for the Georges Bank stock region and was used to calculate the factors-at-length that were used to convert the 2009-2010 Bigelow survey indices to Albatross units for use in VPA model calibration. CVs-at-age for the tuning indices were highest for the Canadian spring surveys (ranging from 0.21 to 0.41), followed by the U.S. fall survey indices (ranging from 0.16 to 0.28) and U.S. spring indices (ranging from 0.13 to 0.24).

Despite considerable inter-annual variability, the NEFSC fall survey relative abundance indices show an increasing trend during the 1970's, followed by a declining trend during the 1980s to a time series low in 1991. Thereafter, relative abundance increased through 2001 then declined to a level below the 1963-2009 median during 2005-2007. In 2009, fall relative abundance reached the second highest point in the time series, but declined drastically in 2010 to a level slightly below the time series median. Trends in the NEFSC spring survey relative abundance indices exhibited more inter-annual variability, but were similar to the fall survey time series after 1982. NEFSC spring survey abundance indices were at record low levels during 2004-2007. The second highest abundance index of the time series occurred in 2008. However, most of the fish were caught at two consecutively sampled stations and relative abundance declined severely the following year and was at the time series median in 2010. Relative abundance trends in the Canadian survey were similar to those in the NEFSC spring survey during most years but were of greater magnitude during blocks of years (1988-1990 and 1993-1997). Similar to relative abundance indices from the NEFSC spring surveys, indices from the Canadian surveys were at the lowest levels observed during 2005-2007 but then declined to well-below the time series median in 2009 and 2010.

Although the survey numbersat-age were highly variable, large cohorts appeared to track through the numbers-at-age matrices, for the NEFSC surveys, for the 1980, 1987, 1994, 1998-2001, and 2006 cohorts. Age truncation occurred between 1983 and 1997, during which time the population was dominated by four age groups rather than seven or more. During 1997-2004, the age structure improved but has since become truncated again. Both the U.S. and Canadian spring surveys show reduced numbers of age 1-3 fish (and age 4 fish in the CA surveys) during 2000-2007. The Canadian spring survey did not show the same magnitudinal increase in ages 1-6 fish that was evident in the NEFSC spring surveys during 2008-2010.

Term of Reference 3: ***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 area-swept biomass estimates. Investigate if implied survey gear or catchability estimates are reasonable. Include a historical retrospective analysis to allow a comparison with previous assessment results.***

The final VPA model differed from the 2008 VPA model because M was increased from 0.2 to 0.3, discards from the Canadian scallop dredge fleet were added to the catch-at-age, and a new maturity schedule was used. Similar to the 2008 GARM assessment results, very mild retrospective patterns were present for terminal years 2001-2009. However, a flip in terminal year estimates of fishing mortality (overestimation during 2006-2009 and underestimation during 2002-2005) and spawning stock biomass (underestimation during 2006-2009 and overestimation during 2002-2005) occurred. There was no retrospective pattern for terminal year age 1 recruitment, but the estimates were highly variable Residuals patterns were evident for a number of ages within each of the three sets of VPA calibration indices.

VPA estimates of survey catchability increased with age for all three surveys. Catchabilities were higher for the NEFSC fall surveys than the NEFSC spring surveys (e.g., q = 0.20 and 0.28 for age 6, respectively), but qs-at-age between the two surveys were not significantly different. Catchabilities for the Canadian spring surveys can be compared across ages but not between surveys because the vessels and gear were different. The catchabilities of ages 1-3 fish were significantly lower than for ages 5-7+ fish.

Fishing mortality rates were highest during 1984-1993, ranging between 0.57 and 1.17, but then declined to levels ranging between 0.31 and 0.51 during 1994-1998. Fishing mortality rates were low (0.26-0.27) during 1999 and 2000, then increased rapidly to 0.85 in 2003, followed by a rapid decline to the second lowest level in the time series (0.20) in 2006. Fishing mortality increased slightly during 2007-2009, but then declined to 0.15 in 2010.

SSB declined rapidly from a time series peak of 17,380 mt in 1982 to 6,256 mt in 1985, then increased slightly through 1987 to 8,082 mt. After 1987, SSB declined again to a time series low of 3,424 mt in 1995. SSB subsequently increased to 13,790 mt in 2000, but then declined to 5,305 mt in 2005. Thereafter, SSB increased and totaled 9,703 mt in 2010.

Trends in age 1 recruitment indicated several periods of rise-and-fall. Recruitment increased from 8.3 million fish in 1983 to a time series peak of 26.3 million fish in 1988, and then declined to 5.2 million fish in 1993. Recruitment increased again to fairly high levels during 1995-1999 (16.2-22.8 million fish) then declined to the second lowest level on record (5.5 million fish) in 2004. Recruitment increased to 18.8 million fish in 2008, but then declined to the lowest level in 2009 (4.0 million fish). Recruitment increased to a very high level (22.5 million fish) in 2010, then declined again in 2011 to near the 2009 level. The 2011 estimate is uncertain because it is based solely on the geometric mean of recruitment during 2003-2009.

Term of Reference 4: ***Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on model performance (in TOR-3).***

The interpretation of TOR4, by the Southern Demersal Working Group (SDWG), was that the variance of the commercial landings due to the area-allocation scheme (for 1995 and onward) should be used as the basis for estimating the magnitude of landings that might be lost or gained for the stock-specific assessments, and that the assessment models should be run with such potential biases incorporated in order to evaluate their effects on estimates of F, SSB, and R.

For the GB winter flounder stock, total landings for 1995-2010 have a calculated Proportional Standard Error (PSE; due to the aforementioned commercial landings area-allocation procedure) ranging from 0.7% to 1.3%. The total discard PSEs during 1995-2010 ranged from 1% to 56%. Because the PSEs for the landings are low, and the landings accounted for 69-94% of the total catch during 1982-2010, the total catch-weighted annual PSEs ranged from 1.2% to 8.2% and averaged 3.9% (unweighted) for the 1982-2010 time series.

The SDWG developed an exercise using the 2008 GARM assessment data and VPA model in an initial response to TOR4 (Terceiro MS 2011) and concluded that the application of a annually varying unidirectional "bias-correction" provides stock size estimates and BRPs that scale either up or down by about the same average magnitude as the landings gain or loss.

Since the initial exercise of SDWG WP3, the SDWG concluded that the calculated variance of the area-allocated commercial landings likely underestimates the true error. More work was done to estimate the error in the commercial landings due to misreporting of commercial landings to statistical area at allocation level “A” reporting level in mandatory Vessel Trip Reports (Palmer and Wigley MS 2011). The perceived under-reporting of statistical areas in the VTR data led to minor (< 5%) differences in the overall species landings allocations. Therefore, the SDWG elected to an additional 5% PSE to the PSE values of the GB total landings during 1995-2010. This increased the 1995-2010 average landings PSE from 0.9% to 5.7%, and increased the average 1982-2010 catch PSE from 4.0% to 6.2%, with a range of 2.7% in 1983 to 13.7% in 2010.

The catch in the final assessment model was increased/decreased by the annually varying catch PSEs and models were re-run to provide an additional measure of the uncertainty in assessment estimates. As noted in SDWG WP3, the application of a annually varying "bias-correction" in one direction in such an exercise provides stock size estimates that scale up or down by about the same average magnitude as the gain or loss. For the final VPA model results, fishing mortality did not change, on average (out to three decimal places), and the range in 2010 F was 0.154 to 0.162. SSB changed by – 1.0% and +7.9%, on average, and the range in 2010 SSB was 9,636 mt - 10,504 mt.

Term of Reference 5: ***Examine the effects of incorporating environmental factors in models of population dynamics (e.g., spring water temperatures in an environmentally-explicit stock recruitment function).***

The conclusion from the analysis was that recruitment in the coastal stocks of winter flounder (GOM and SNE-MA) were linked to air temperatures during winter, when spawning occurs, but there was no evidence for an air temperature effect on recruitment in the Georges Bank stock; the environmentally-explicit models (which also included a Gulf Stream index) did not provide a better fit compared to the standard stock recruitment model. The Georges Bank stock experiences water temperatures that are affected by both local air temperatures and more importantly, large-scale advective supply of relative cold, fresh water associated with the Labrador Current. Examining other environmental variables which may affect recruitment in the Georges Bank stock (e.g., hydrographic circulation patterns on Georges Bank in relation to larval abundance) is listed below as a future research recommendation.

Term of Reference 6: ***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}, $B_{THRESHOLD}$, and F_{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 specification of FMSY and BMSY-based reference points relies on a stock-recruitment relationship. As a result, the 2008 GARM Biological Reference Point Review Panel (NEFSC 2008) concluded that MSY-based BRPs should be adopted when the stock-recruitment relationship is informative, and if not, then the Panel recommended the use of F40%MSP as a proxy for FMSY, similar to the previous recommendation from a separate BRP Working Group for many of the groundfish stocks (NEFSC 2002a), and a BMSY proxy computed using a non-parametric, empirical approach.

For Georges Bank winter flounder, the 2008 GARM BRP Review Panel concluded that the Beverton-Holt stock-recruitment model (Beverton and Holt 1957) fit, using data from a VPA, did not provide feasible results either without a prior (h=1) or with a prior (the fit was highly dependent on the assumed prior on unfished recruitment, R_0). Thus, the Panel recommended that the non-parametric empirical approach be used to estimate biological reference points based on: 1) the final VPA model results, 2) the estimate of F40%MSP as a proxy for F_{MSY} (derived using the most recent five-year average of fishery selectivity and weights-at-age and the maturity-at-age time series average), and 3) a long-term (100 year) stochastic projection using the cumulative distribution function of observed recruitment (1983-2007 recruitment at age 1, the 1982-2006 year classes) to estimate MSY and SSBMSY. The existing biomass target is SSBMSY at 40% MSP (= 16,000 mt) and the minimum biomass threshold is 50% of the target (= 8,000 mt). The fishing mortality threshold is an F40%MSP (= 0.26). Amendment 16 defines the fishing mortality target as the mortality associated with the Annual Catch Limit (ACL).

Two sets of candidate BRPs (i.e., FMSY and SSBMSY versus F40% and SSB40%) were brought forward from the current assessment for review by the SARC because the SDWG could not reach consensus on whether the stock-recruit relationship from the Beverton-Holt model was informative, and consequently, whether FMSY was well-estimated. However, the SARC had concerns about the prior on unfished steepness and the fact that the stock-recruitment data for the Georges Bank stock was less informative than the SNE/MA data for predicting recruitment at low spawner levels, making direct estimation of the spawner-recruit relationship difficult without external information. The SARC also concluded that steepness values should be similar between winter flounder stocks. Therefore, the steepness log-likelihood profiles of the two stocks were used in selecting fixed values for steepness with which to estimate FMSY for each stock. Fixed steepness values of 0.61 and 0.78 were selected for the SNE/MA and Georges Bank stocks, respectively. Precision estimates for the resulting FMSY reference point estimates were not possible due to the fixed for steepness. Candidate BRPs estimated for the Georges Bank winter flounder stock which were used to determine 2010 stock status were: FMSY (Fthreshold) = 0.42; SSBMSY (Btarget) = 11,800 mt; ½ SSBMSY (Bthreshold) = 5,900 mt and MSY = 4,400 mt.

Term of Reference 7: *Evaluate stock status (overfished and overfishing) with respect to the "new" BRPs (from TOR 6), and with respect to the existing BRPs (from a previous accepted peer review) whose values have been updated.*

During 2010, the Georges Bank winter flounder stock was not overfished and overfishing was not occurring. The fishing mortality rate in 2010 (= 0.15) was below FMSY (= 0.42) and spawning stock biomass in 2010 (= 9,703 mt) was above the SSB_{MSY} threshold (= 5,900 mt). In the current assessment, the assumed value for M was increased from 0.2 to 0.3. As a result, the SDWG concluded that a comparison of the 2010 F and SSB estimates from the current assessment with the existing reference points (estimated assuming an M of 0.2) was not appropriate.

Term of Reference 8: ***Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs) under a set of alternative harvest scenarios. If the stock needs to be rebuilt, take that into account in these projections.***

a. Provide numerical short-term projections (3-5 yrs, or through the end of the rebuilding period, as appropriate). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment (e.g., terminal year abundance, variability in recruitment). Take into consideration uncertainties in the assessment and the species biology to describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming or remaining overfished, and how this could affect the choice of ABC.

Stochastic medium-term projections of future stock status, during 2011-2017, were made based on the current assessment results for the final VPA model and of the final set of candidate BRPs. Maturity-at-age, and mean weights and fishery selectivity patterns-at-age, estimated for the most recent 5 years in the assessment (2006-2010), were used to reflect current conditions in the stock and fishery. The projections assumed that a catch of 2,118 mt (for the FMP Framework 44 fishing year beginning May 1) would be landed as the calendar year catch in 2011. The projections incorporated uncertainty in the current population estimate, via bootstrap replicates, and variability in predicted recruitment. A parametric Beverton-Holt model with log-normal error was used and recruitment variability was generated by randomly sampling from the estimated error distribution of the fitted stock–recruitment model.

The regulations require rebuilding of the Georges Bank stock, with at least 75% probability, by 2017. The projections indicated that rebuilding to SSBMSY (= 11,800 mt) is expected to be achieved with 78% probability in 2012 when fishing at 75% of FMSY (=0.315) with a catch of 2,118 mt in 2011.

Vulnerability, productivity and susceptibility of the Georges Bank winter flounder stock using several methods. Uncertainty was evaluated using model estimates of precision and qualification of other uncertainties. The age-based VPA model and associated MSY reference point evaluations provide a relatively comprehensive and synthetic evaluation of vulnerability that is entirely consistent with stock status determination and projection. With respect to status determination, vulnerability and susceptibility were accounted for with regards to estimation of F in 2010, but precision estimates for F_{MSY} were not possible due to the use of a fixed steepness value in the Beverton-Holt stock-recruit model. Stock vulnerability and susceptibility were also accounted for in the stock rebuilding projection. All components of productivity (reproduction, individual growth, and survival) were also explicitly accounted for in stock status determination and projections. Reproduction was monitored as age-1 recruitment, and projected as a function of SSB (the product of abundance, weight- and maturity-at- age). Individual growth was monitored as empirical size at age, and projected as recent mean size at age. Survival was accounted for based on model estimates of fishing mortality and selectivity as well as assumed natural mortality, which was informed by tagging analysis.

Uncertainties that were not accounted for by assessment and reference point models were evaluated using model diagnostics. Standard model diagnostics (e.g., residual analyses, retrospective analyses) were used for model validation. Vulnerabilities that were not accounted for by assessment and reference point models were evaluated using exploratory modeling, habitat observations and testing the influence of environmental factors on recruitment dynamics. A small portion of the stock (5-17% of the total catch during 2004-2010) is not regulated by the US, yet is susceptible to fishing (i.e., incidental catches) by the Canadian scallop dredge and groundfish bottom trawl fleets. Winter flounder discards in the latter fleet are unknown.

b. Develop plausible hypotheses (e.g., mixing among the three stocks) which might explain any conflicting trends in the data and undertake scenario analyses to evaluate the consequences of these alternate hypotheses on ABC determination.

The SDWG has initiated research pursuing the use of a more complex model (i.e., Stock Synthesis) to maintain separate fishery and survey catch for the three current stock units, while allowing a small amount (a few percent) of exchange between the stock units based on information from historical tagging. However, development of that research has not progressed sufficiently to be made available for peer review at this time.

Term of Reference 9: ***Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.***

A list of progress made on research recommendations from prior assessments and a prioritized list of new research recommendations that would improve the assessment of the Georges Bank winter flounder stock is presented below under TOR 9.

Terms of Reference

1. Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data.

2. Present survey data being considered and/or used in the assessment (e.g., regional indices of abundance, recruitment, state and other surveys, age-length data, etc.). Characterize uncertainty in these sources of data.

3. 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 area-swept biomass estimates. Investigate if implied survey gear or catchability estimates are reasonable. Include a historical retrospective analysis to allow a comparison with previous assessment results.

4. Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on

model performance (in TOR-3).

5. Examine the effects of incorporating environmental factors in models of population dynamics (e.g., spring water temperatures in an environmentally-explicit stock recruitment function).

6. 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}, $B_{THRESHOLD}$, and F_{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.

7. Evaluate stock status (overfished and overfishing) with respect to the "new" BRPs (from TOR 6), and with respect to the existing BRPs (from a previous accepted peer review) whose values have been updated.

8. Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs) under a set of alternative harvest scenarios. If the stock needs to be rebuilt, take that into account in these projections.
   a. Provide numerical short-term projections (3-5 yrs, or through the end of the rebuilding period, as appropriate). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment (e.g., terminal year abundance, variability in recruitment).
   b. Take into consideration uncertainties in the assessment and the species biology to describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming or remaining overfished, and how this could affect the choice of ABC.
   c. Develop plausible hypotheses (e.g., mixing among the three stocks) which might explain any conflicting trends in the data and undertake scenario analyses to evaluate the consequences of these alternate hypotheses on ABC determination.

9. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.

Assessment history

The Georges Bank winter flounder stock was assessed in November, 2001 at SAW/SARC 34 (NEFSC 2002b). The assessment results and biological reference points (BRPs) were based on an ASPIC biomass dynamics model (Prager 1995) which incorporated landings (1964-2000) and biomass indices from the NEFSC autumn (1963-2000) and spring (1968-2001) bottom trawl surveys. Model results indicated a reasonable fit to the input data and that yield has been below surplus production since 1994. Relative estimates of mean biomass (B_t/B_{MSY}) declined sharply during 1977-1994, then increased to B_{MSY} in 2000. Relative fishing mortality rates (F_t/F_{MSY}) were at or below F_{MSY} during1994-2000. During 2000, the stock was not overfished and overfishing was not occurring.

During the 2002 GARM (NEFSC 2002c), stock status was assessed from the results of an updated run of the SAW 34 ASPIC model formulation. Data included in the model were: the NEFSC survey biomass indices for autumn of 1963-2001 and spring of 1968-2002, and total landings during 1964-2001. Fishing mortality rates declined sharply during 1993 and 1999, from 0.71 to 0.14, and were at or below F_{MSY} (= 0.32) during 1995-2001. Average total biomass increased after 1994 and was slightly above B_{MSY} during 2001. There was no retrospective pattern in the ASPIC-derived estimates of fishing mortality rates or total biomass. The biological reference point estimates from the SARC 34 ASPIC model were also recommended for implementation by the 2002 Working Group on Re-estimation of Biological Reference Points for New England Groundfish (NEFSC 2002a). The existing reference points were: F_{MSY} = 0.32, B_{MSY} = 9,400 mt, and MSY = 3,000 mt. The 2002 Working Group concluded that the use of absolute reference point values from the ASPIC model (based on total biomass rather than exploitable biomass) were appropriate because the NEFSC surveys appeared to measure the biomass of the exploitable portion of the stock. The 2001 fishing mortality rate estimate was 0.25 and the 2001 total biomass estimate was 9,805 mt. Therefore, the stock was not overfished and overfishing was not occurring in 2001.

The stock was assessed next in September 2005 during GARM 2 (NEFSC 2005). The assessment consisted of an updated run of the SARC 34 ASPIC model (Prager 2004) formulation. Input data to the model included landings (1964-2004) and NEFSC fall (1964-2004) and spring (1968-2005, lagged back one year) survey relative biomass indices. ASPIC-based biological reference points are re-estimated each time the model is run and model estimates of relative total biomass (B_t/B_{MSY}) and fishing mortality rates (F_t/F_{MSY}) are more precisely estimated than the absolute values (Prager 1995). Therefore, the 2005 GARM review panel concluded that bias-corrected relative estimates of annual total biomass and fishing mortality rates from the updated ASPIC model run should be compared to relative biological reference points (biomass threshold = 0.5, fishing mortality rate threshold = 1.0) to determine stock status. In 2005, the stock was not overfished, but overfishing was occurring.

For the 2008 GARM (NEFSC 2008), a VPA model was used because of improved biological sampling of the fishery since SARC34, the need to assess changes in the population's truncated age structure, and to avoid the pitfalls associated with the biomass-based ASPIC model. Model inpuat data included: catch - at-age data for ages 1-7+, including initial estimates of discards-at-age, for the U.S. bottom trawl and scallop dredge fleets and U.S. and Canadian landings. At the GARM 3 BRP meeting, the review panel determined the stock-recruitment relationship predicted from a Beverton-Holt model was not reliable. As a result, BRPs were derived based on the empirical cumulative distribution function of age 1 recruitment, for 1982-2007, from the VPA model. A 100-year, stochastic projection was run using an age-structured projection model and assuming a constant harvest scenario of $F_{40\%}$ = 0.26 (estimated from a per-recruit model) to predict the median MSY40% (= 3,500 mt) and SSB40% (=16,000 mt) under equilibrium conditions. The 2007 fishing mortality rate (= 0.28) was above the F_{MSY} proxy (= 0.26), indicating that overfishing was occurring in 2007. The spawning stock biomass in 2007 (= 4,964 mt) was well below the SSB_{MSY} target (8,000 mt), indicating that the stock was also overfished in 2007. The 2007 estimates of average F and SSB did not require adjustments for the mild VPA retrospective pattern because the 2000-2006 average Mohn's rho values for average F and SSB were within the 80% confidence limits of the average F and SSB estimates.

The current assessment is an update of the VPA model formulation from the 2008 GARM (NEFSC 2008), including data for 1982-2010, but with the addition of discards-at-age for the Canadian scallop dredge fleet, an assumed increase in M from 0.2 to 0.3, and a new maturity schedule.

Growth and maturity

Winter flounder in the Gulf of Maine and Southern New England reach a maximum size of around 2.25 kg (5 pounds) and 60 cm. On Georges Bank fish may reach a maximum length of 63.5 cm and weight up to 3.6 kg (8 pounds; Bigelow and Schroeder 1953). An updated compilation and analysis of the NEFSC and MADMF survey growth and maturity data for 1976-2010 indicated the following maximum age and length and von Bertalanffy growth parameters that generally support the current stock identifications (Figure B1):

GOM: N = 16,010 fish, maximum age = 15 (55 cm); maximum length = 61 cm;
L_∞ = 46.4 cm, k = 0.2727

GBK: N = 6,311 fish, maximum age = 18 (50 cm), maximum length = 70 cm;
L_∞ = 57.9 cm, k = 0.2829

SNE: N = 23,593 fish, maximum age = 16 (51 cm), maximum length = 60 cm;
L_∞ = 46.5 cm, k = 0.3184

Previous assessments of SNE-MA winter flounder (NEFSC 1999; NEFSC 2003; NEFSC 2005 and NEFSC 2008) have included maturity schedules derived using data for females from the MA DMF spring surveys and published in O'Brien et al. (1993), who fit probit regression models assuming lognormal error to the maturity at age data to estimate the proportions mature at age.

In response to a SAW 28 research recommendation (NEFSC 1999), the 2002 SAW 36 (NEFSC 2003) examined NEFSC spring trawl survey data for the 1981-2001 period in an attempt to better characterize the maturity characteristics of the SNE/MA winter flounder stock complex. The NEFSC maturity data indicated earlier maturity than the MADMF data, with L50% values ranging from 22-25 cm, rather than from 28-29 cm, and with ~50% maturity for age 2 fish, rather than ~50% maturity for age 3 fish (NEFSC 2003). This trend was confirmed through histological analyses by McBride et al. (MS 2011) which indicated that age 2 fish are likely not mature (also see SDWG WP 13). Therefore, given the results from the SAW 36 comparisons and the histological study, the SDWG concluded that the MADMF spring survey data continue to provide the best macroscopic evaluation of the maturity stages for SNE/MA winter flounder and that 1982-2008 time series of maturity estimates at age should be used for SARC 52 assessment.

Georges Bank winter flounder spawn during March-May, with a peak in April (Smith 1985).The maturity schedule used in the VPA model during the previous stock assessment (NEFSC 2008), shown in the following table, was the time series average during 1982-2007 for females caught during NEFSC spring surveys (which generally occur on Georges Bank during April). Probit regression models assuming lognormal error were fit to the maturity at age data to estimate the proportions mature at age. Given the finding that the NEFSC spring surveys suggest that the age at 50% maturity occurs one year earlier than the A50 computed from the MA DMF surveys and the histological results in the McBride et al. (2010), the SDWG adopted the maturity schedule shown in the table below. The maturity schedule was estimated as a 3-year moving window based on an adjustment of the female maturity-at-age data from the 1981-2010 NEFSC spring surveys (strata 13-23). Based on the female maturity at length data for the 57 Georges Bank fish from the histology study (Figure B2), fish > 30 cm TL were misidentified macroscopically, at sea, as immature fish and fish < 38 cm were misidentified as resting fish. Therefore, immature fish > 30 cm (= 7% of the immature fish during 1981-2010) and resting fish < 38 cm (= 28% of the mature fish during 1981-2010) were deleted prior to fitting the probit regression model. All of the deleted fish were ages 2 and 3. The resulting female A50 values and their 95% confidence intervals are shown in Figure B3.

Age	1	2	3	4	5	6	7+
Stock, assessment period (years included)							
GB, 2008 GARM (1982-2007)	0.08	0.54	0.94	1.00	1.00	1.00	1.00
SNE/MA, current assessment (1982-2008)	0.00	0.08	0.56	0.95	1.00	1.00	1.00
GB, current assessment (1981-2010)	0.00	0.09	0.90	1.00	1.00	1.00	1.00

Instantaneous natural mortality (M)

The SDWG adopted a change in the instantaneous rate of natural mortality (M) for the winter flounder stocks. The value of M previously used in all assessments was 0.2 for all ages and years, and was based on the ICES 3/Tmax "rule-of-thumb" (e.g., see Vetter 1998 and Quinn and Deriso 1999)using observed maximum ages for winter flounder (Tmax) of about 15. The current observed Tmax values for the three stock units are GOM = 15 years, GBK = 18 years, and SNE/MA = 16 years (see previous Growth and Maturity section). The adopted change increases this rate to 0.3 for all stocks, ages and years. Evidence can be found in the literature and current model diagnostics to support the increase.

Literature values of M from tagging studies and life history equations indicate M for winter flounder is likely higher than 0.2. Dickie and McCraken (1955) carried out a tagging study in St. Mary Bay, Nova Scotia, Canada (GOM Stock) and estimated a percentage natural mortality rate to be 30% (M = 0.36). Saila et al. (1965) applied Ricker's equilibrium yield equation to winter flounder from Rhode Island waters (Tmax = 12) and using F values from Berry et al. (1965) calculated M to be 0.36. Poole (1969) analyzed tagging data from New York waters from five different years and estimated values for M of 0.54 (1937), 0.33 (1938), 0.5 (1964), 0.52 (1965), and 0.52 (1966). Finally, an analysis of tagging data from a large scale study along the coast of Massachusetts provided a percentage natural mortality rate of 27%, or M = 0.32 (Howe and Coates 1975). For this assessment, a re-analysis of the Howe and Coates (1975) tagging data was conducted using a contemporary tagging model to estimate natural mortality (Wood WP 15). The tagging model fit to the data was the instantaneous rates formulation of the Brownie et al. (1985) recovery model (Hoenig et al. 1998). This work provided an M of 0.30 with 95% confidence interval from 0.259 to 0.346.

Values derived from life history equations found in the literature also support a higher estimate of M for winter flounder. Three equations were used along with a maximum age (Tmax) of 16 to derive estimates of M equal to 0.28, 0.26, and 0.19 (the equations from Hoenig 1983, Hewitt and Hoenig 2005, and ICES, respectively). A newly proposed method from Gislason et al. (2010), based on SNE/MA stock mean size at age (Ages 1-16) and von Bertalanffy growth parameters, estimated M to be 0.37 (see text table below).

Values of natural mortality (M) for winter flounder found in the literature and derived using life-history equations.

Study	**Method**	**M**

ICES rule-of-thumb	Equation: 3/Tmax	0.19
Hewitt and Hoenig 2005	Equation: 4.22/Tmax	0.26
Hoenig 1983	Equation: 1.44-0.982*ln(Tmax)	0.28
Howe and Coates 1975	Analysis of Tagging Data	0.32
Wood 2011 (WP15)	Re-analysis of Howe and Coates 1975	0.30
Poole 1969	Analysis of Tagging Data from 1938	0.33
Dickie and McCraken 1955	Analysis of Tagging Data	0.36
Saila et al. 1965	Ricker Equil. Yield Equation and Tmax	0.36
Gislason et al. 2010	Equation: Mean size at age and VBG	0.37
Poole 1969	Analysis of Tagging Data from 1964	0.50
Poole 1969	Analysis of Tagging Data from 1965	0.52
Poole 1969	Analysis of Tagging Data from 1966	0.52
Poole 1969	Analysis of Tagging Data from 1937	0.54

Preliminary assessment population model run diagnostics also, in general, support a higher value for M. Profiles in mean squared residual for ADAPT VPA SNE/MA stock models indicate best fits for M in the range of 0.2 to 0.3. The likelihood profile of initial ASAP SCAA model runs for the SNE/MA stock indicates a best fit for M= 0.6. Model runs from Rademeyer and Butterworth SCAA (ASPM) model (2011) at M equal to 0.2, 0.3, and 0.4 also revealed decreasing negative log-likelihood as M is increased for GOM and SNE/MA stock models (see the following two text tables).

Results of SCAA for the Gulf of Maine winter flounder for each combination of 3 levels of natural mortality (*M*=0.2, 0.3 and 0.4, constant throughout the assessment period) and 3 weightings of the survey CAA likelihood (*w*=0.1, 0.3 and 0.5). The runs with w=0.3 and 0.5 have both commercial and survey selectivities flat at older ages, while the runs with w=0.1 have only the commercial selectivity flat. Displayed values are the negative log-likelihoods of each model.

		M	
Weighting	0.2	0.3	0.4
0.1	-123.2	-126.6	-129.1
0.3	-156.9	-177.2	-196.1
0.5	-255.6	-263.2	-280.8

Results of SCAA for the SNE/MA winter flounder for 3 levels of natural mortality for Base Case 2. Displayed values are the negative log-likelihoods of each model.

	M		
	0.2	0.3	0.4
-LL	-123.2	-126.6	-129.1

The SDWG also considered other evidence that might justify an increase in M for winter flounder. The NEFSC's food habits database (Smith and Link 2010) was examined to identify the major fish predators of winter flounder. These predators include Atlantic cod, sea raven, monkfish (goosefish), spiny dogfish, winter skate and little skate. A preliminary examination was undertaken to determine the prominence of winter flounder in the diets of these predators, across all seasons, years, size classes of predator, sizes of prey, and geographic locales. The overall frequency of occurrence of winter flounder in the stomachs is not a common or high occurrence (see text table below), always less than 0.15%.

Occurrence of winter flounder in their major fish predators

	Number of stomachs	Occurrences of winter flounder	% frequency of occurrence
Spiny dogfish	67,565	27	0.040%
Winter skate	17,708	6	0.034%
Little skate	28,725	6	0.021%
Atlantic cod	20,142	27	0.134%
Sea raven	7,968	10	0.126%
Goosefish	10,742	12	0.112%

Further, the contribution of winter flounder to the diets of these predators species is also notably small (see text table below), usually less than 0.4%.

Contribution of winter flounder to the diets of their major fish predators

	% diet composition, by weight	95% CI
Spiny dogfish	0.2049%	0.10678
Winter skate	0.1454%	0.16008
Little skate	0.0124%	0.01618
Atlantic cod	0.3172%	0.24032
Sea raven	0.8831%	0.78407
Goosefish	0.2492%	0.25947

Understandably, the temptation exists to evaluate these relatively low diet contributions with respect to consumptive removals of winter flounder as compared to winter flounder stock abundance and (relatively low) landings, initially using *ad hoc* or proxy methods. However, just as one would not do so when assessing the status of a stock without a fuller exploration of all the sensitivities, uncertainties and caveats of the appropriate estimators and parameters, the SDWG did not recommend doing so for scoping winter flounder predatory removals at this time. The SDWG also noted that for percentages as low as those observed, when allocated to the three winter flounder stocks and explored seasonally or as a time series, there are going to be large numbers of zeroes and attendant uncertainties and variances that would logically offset any potentially high individual predator total population-level consumption rates. Thus, the SDWG does not provide comment as to the merit of exploring or relative magnitude of the issue, but recommends that the topic should be forwarded as an important research recommendation.

Other sources of increased natural mortality may come from perceived increases in seal populations along the New England coast, which are known to be predators of winter flounder (Ampela 2009). Population size was estimated at 5,600 seals in 1999 (Waring et al. 2009) and a current survey is being conducted to estimate the size of the seal population. However, no time series of seal abundance or consumption of winter flounder is available.

<u>Additional analyses conducted during the SARC</u>

For Georges Bank winter flounder, two sets of candidate BRPs (i.e., FMSY and SSBMSY versus F40% and SSB40%) were brought forward from the current assessment for review by the SARC because the SDWG could not reach consensus on whether the the stock-recruit relationship from the Beverton-Holt model was informative, and consequently, whether FMSY was well-estimated. However, the SARC did not select either set of BRPs. Rather, the SARC concluded that the estimation of a stock-recruit relationship for the Georges Bank stock was difficult without external information and that the use of a steepness prior for Pleuronectids based on Myers et al (1999) was inappropriate. The SARC also concluded that steepness should be similar across all three winter flounder stocks. Therefore, given that the SNE/MA stock-recruit relationship was more informative, the SARC used the log-likelihood steepness profiles of each stock to select a fixed steepness value with which to rerun the Beverton-Holt model to obtain a final estimate of FMSY. The methods and results of the analyses are discussed below under TOR 6.

Term of Reference 1: ***Estimate catch from all sources including landings and discards. Characterize the uncertainty in these sources of data.***

Landings

Statistical Areas used for reporting fishery data for the Georges Bank winter flounder stock include: 522-525, 542, 551-552, and 561-562 (Figure B4). Several different methods have been used to collect the landings, fishing area and effort data. During 1963 through April of 1994, U.S. commercial landings, effort, fishing area, and other fishery-related data were collected and entered into Northeast Region Commercial Fisheries Database (CFDBS) by NMFS port agents, who entered landings data from all dealer purchase receipts and interviewed a subset of captains to obtain information about fishing location and effort (Burns *et al*. 1983). During May of 1994-2003, reporting of landings and other associated trip data was mandatory for dealers issued federal permits to purchase groundfish. The data were collected and entered into the CFDBS by NMFS port agents. Since 2004, the landings and associated trip data have been self-reported, electronically, by federally permitted dealers. Beginning in May of 1994, mandatory reporting of fishing location (Statistical Area) and effort data, gear type, estimated catch, and other trip-based fishing data were self-reported by fishermen on logbooks (i.e., Vessel Trip Reports or VTRs) and the data were entered into the Vessel Trip Report Database. In order to integrate data from the VTR Database with data from the CFDBS, an "allocation" database was created using a trip-based allocation scheme (Wigley et al. 2008a). Landings data are assumed known and originate from the CFDBS. The allocation determines the area fished and effort information reported on the VTR data and joins this information with the landings data from each trip as reported in the CFDBS. Two levels (A and B) represent vessel-oriented data and two levels (C and D) represent fleet-oriented data. Level A comprises audited VTR trips that have not been grouped and for which a one-to-one match exists between the VTR and CFDBS fields which define a trip (i.e., year, month, day and permit). Level B comprises VTR trips from Level A that have been pooled by vessel permit, gear group, main species group, and month. Level C comprises VTR trips from Level A that have been pooled by ton class, port group, gear group, main species group, and calendar quarter. Level D comprises VTR trips from Level A that have been grouped by port group. If a CFDBS trip has a corresponding one-to-one match with a VTR trip, then the area fished and the effort information, if present, is transferred directly onto the CFDBS trip record. "A" level trips correspond to pre-1994 trips for which similar information was obtained from a vessel captain via a port agent interview.

During 1995-2010, 63-78% of the commercial landings were allocated to Statistical Area based on a 1:1 match of trips in the Dealer and Vessel Trip Report Databases ("A" level data), with the majority of the remaining trips allocated at the "B" level for which stratification is based on vessel, month, gear group, and species group basis (Table B1). For the Georges Bank winter flounder landings, the Proportional Standard Error (PSE, reported as a %) due to the allocation of landings to Statistical Area, using Vessel Trip Reports, ranged between 0.7 and 1.3% during 1995-2010 (Table B2).

There are no significant recreational landings of winter flounder from Georges Bank. Commercial landings data were available for 1964-2010. Since 1964, total landings have been predominately from the U.S groundfish trawl fishery, but landings have also been reported for the Canadian groundfish trawl fisheries, as bycatch in the haddock and cod fisheries (Heath Stone pers. comm.). During 1965-1977, landings were also reported by the former USSR; reaching a peak of 1,699 mt in 1972 (Table B3, Figure B5). Canadian landings generally comprised a low percentage (1-2 %) of the total landings until 1994, at which time Canadian landings increased rapidly from 6 % of the total to a peak of 24 % in 2001 (529 mt). The increasing trend in Canadian landings occurred primarily during the second half of the year because since 1994 Canadian groundfish fisheries on Georges Bank have, for the most part, been closed during January-May (Van Eeckhaute and Brodziak 2005). After 2001, Canadian landings declined rapidly to 1.5% in 2007 (12 mt). During 2008-2010, Canadian landings were very low, comprising only 1-3% of the total landings.

Total landings increased during 1964-1972, reaching a peak of 4,509 mt in 1972, then declined to1,892 mt in 1976 (Figure B5, Table B3). A sustained period of high landings occurred during 1977-1984, ranging from 3,061-4,009 mt. After 1984, landings gradually declined to the lowest level in the time series, 783 mt in 1995, but then increased again to 3,139 mt in 2003. Thereafter, landings declined rapidly and reached the second lowest level on record in 2007 (807 mt). During the time period included in the stock assessment model, 1982-2010, total landings averaged 1,950 mt and were slightly below this average in 2009 (1,670 mt) and 2010 (1,297 mt).

Most of the U.S. landings (92-100%) are taken with bottom trawls and most of the remainder is taken by the scallop dredge fleet (Table B4). During most years since 1982, landings taken by the scallop dredge fleet have been less than 1% of the U.S. total. However, a high period of landings by the scallop dredge fleet (4-8% of the total landings) occurred during 1988-1993 and in 2005-2006 (6% and 3%, respectively, of the total landings).

The spatial distribution of winter flounder landings on Georges Bank has largely been affected by complex management regulations. During 1982-1993, prior to the implementation of groundfish Closed Areas I and II (Figure B6), most of the Georges Bank landings of winter flounder were taken in the two northern SAs, 522 and 562. Since 1994, portions of the four SAs where most of the landings occur (522, 525, 561 and 562) have been closed, for the most part, to groundfish bottom trawl fishing (Figure B6). During 1994-2001, most of the landings occurred in SA 522 (37-69%), but then shifted to SA 562 during 2002-2005, where 38-54% of the landings occurred (Figure B7). With implementation of the Eastern (SAs 561 and 562) and Western US/CA Areas (SAs 522 and 525) in May of 2004, which was linked to the establishment of total allowable catches (TACs) for cod, haddock and yellowtail for the US versus CA within their respective EEZs, landings began increasing again in SAs 522 and 525. The shift in where the predominant landings occurred (from the Eastern to the Western U.S./CA Area), after 2004, may have been attributable, in part, to the 2005 requirement to use a haddock separator trawl when fishing in the Eastern U.S./CA Area as well as closures of this Area when cod, haddock or yellowtail quotas are reached. The haddock separator trawl was designed to catch haddock but to reduce incidental catches of other demersal finfish species. During 2006-2009, most of the landings (42-53%) were again taken in SA 522 with most of the remainder taken in SA 525. In 2010, 41% and 38% of the landings were taken in SA 522 and SA 525, respectively (Figure B7).

Discards in U.S. fisheries

Estimates of Georges Bank winter flounder discards in U.S. fisheries, during 1964-2010, are provided for the large mesh bottom trawl fleet (codend mesh size ≥ 5.5 inches), small mesh groundfish fleet (codend mesh size < 5.5 inches), and the sea scallop dredge fleet ("limited permits" only) in Table B5. Discards (mt) from each of the three fleets, during 1989-2010, were estimated based on fisheries observer data (obtained the Northeast Fisheries Observer Program Database or NEFOP Database) and the landings data (obtained from the CFDBS) using the combined ratio method described in Wigley et al. (2008b). The 2007 discard estimate from the 2008 GARM Report (NEFSC 2008) was updated. The discard ratio estimator consisted of discards of GB winter flounder divided by the sum of all species kept by a particular fleet and was derived with data from the NEFOP Database. Trip discards ratios were then raised to the level of total landings of all kept species from each trip to compute a total discard estimate for each trip. Discards were estimated by quarter and cells with fewer than one trip were imputed using annual values.

Due to a lack of fisheries observer data, prior to 1989 for the trawl fleets and prior to 1992 for the scallop fleet, discard estimates were hindcast back to 1964 based on the following equation:

$$(1) \qquad \hat{D}_{t,h} = \bar{r}_{c,2003-2004,h} * K_{t,h}$$

where:

$\hat{D}_{t,h}$ is the annual discarded pounds of GB winter flounder for fleet h in year t

$\bar{r}_{c,2003-2004,h}$ is an average combined D/K ratio (discarded pounds of GB winter flounder / total pounds of all species kept) for the fleet h during either 2003-2004 (for the trawl fleets) or 1992-1998 (for the scallop dredge fleet)

$K_{t,h}$ is the total pounds of all species kept (landed) for fleet h in year t

U.S. discards of Georges Bank winter flounder were much higher during 1964-1991 (average = 195 mt) than during 1992-2010 (average = 65 mt). During 1964-1975, U.S. discards were predominately (49-87%) attributable to the large mesh groundfish trawl fleet (listed in Table B6 as the small mesh fleet because the minimum codend mesh size prior to 1982 was less than 5.5 in.), but were primarily attributable to the scallop dredge fleet thereafter. Total U.S. discards, primarily from the scallop dredge fleet, were highest during 1976-1991 (ranging between 142 mt and 348 mt), but then declined to a very low level in 1992 (Table B5, Figure B8). This trend is not attributable to the hindcast discard estimation method used for this time period, but rather the trend in fishing effort (days fished) for the U.S. scallop dredge fleet (NEFSC 2010, Figure B9). After 1991, discards were lower and the trend continued to track the trend in scallop fleet fishing effort. During 1992-2003 discards were low, between 9 and 85 mt, but discards increased thereafter, reaching 188 mt in 2007. The spike in discards during 2010 was primarily attributable to the small mesh fleet for which several high discard ratios were observed on several silver hake trips that occurred on Cultivator Shoals. However, the precision of the 2010 U.S.

discard estimate was low (CV = 0.44). Precision of the annual discard estimates varied by fleet and was generally highest for the large-mesh bottom trawl fleet and lowest for the small mesh bottom trawl fleet, with intermediate values for the scallop dredge fleet (Table B5). During most years since 2005, when trip sampling rates increased substantially in the scallop dredge and large-mesh bottom trawl fleets, precision of the annual discard estimates greatly improved, ranging between 0.09 and 0.44 (Table B5).

Discards in Canadian fisheries

Initial estimates of Georges Bank winter flounder discards in the Canadian scallop dredge fleet were included in the stock assessment. The Canadian sea scallop fishery operating on Georges Bank closes when the annual TAC is caught. There are two sea scallop management areas on Georges (based on depth and productivity) with different TAC's. Landing of groundfish bycatch in the sea scallop fishery has been prohibited since 1996, so presumably all winter flounder bycatch in this fishery is discarded. However, observer coverage was very low and consisted of one trip per month during 2001-July of 2007 and two trips per month thereafter. Observer discards of winter flounder in Canadian sea scallop trips was only available for September 2004-December 2010 and was estimated by staff from the CA Division of Fisheries and Oceans (DFO) using the method of Garvaris et al. (2007). The 2004-2010 average of the proportions of Georges Bank winter flounder discards to sea scallop landings in the Canadian scallop fleet (0.029) was multiplied by the sea scallop landings in the Canadian scallop fleet (CSAS 2010; J. Sameoto 2011 pers. comm.) in order to obtain hindcast winter flounder discard estimates for 1982-2003.

Winter flounder discards in the Canadian sea scallop fishery averaged 123 mt during 2004-2010 and ranged from 44 mt to 252 mt (Table B3). Hindcast discard estimates for the fleet during 1982-2003 ranged between 58 and 199 mt. The associated precision of the estimates is unknown.

Estimates of winter flounder discards in the Canadian bottom trawl fisheries were not available from the CA DFO. Since most of the Canadian landings of Georges Bank winter flounder occur as bycatch in bottom trawl fisheries targeting haddock and cod in (H. Stone pers. comm.), presumably some winter flounder discards also occur in these, and possibly other, groundfish bottom trawl fisheries that operate on Georges Bank. Since the mid-1980's, discarding of groundfish in the Canadian groundfish fisheries on Georges Bank (NAFO Division 5Zj) has been prohibited. However, although there is no discarding of groundfish during observed trips, observer coverage of the groundfish bottom trawl fleet is very low and there is no doubt that discarding of winter flounder occurs because discards for species that are more highly sought after in the Georges Bank Canadian groundfish fisheries (e.g., cod, haddock and yellowtail flounder) have been estimated (Gavaris *et al.* 2010).

Another factor that may also have affected winter flounder discarding in Canadian groundfish trawl fisheries are seasonal closures and gear modifications in the haddock fishery to reduce cod bycatch. Since 1994, the Canadian groundfish fishery on Georges Bank has, for the most part, been subject to a seasonal closure during January 1-June 1. Since 2001-2003, mobile gear vessels without at-sea observers have been required to use separator panels to minimize the bycatch of cod when fishing haddock. This gear modification may also have reduce the bycatch of winter flounder in the haddock fishery because the lower panel has an open cod end to allow cod (and possibly flatfish) to escape, while the upper panel captures and retains haddock. The Canadian yellowtail flounder fishery is required to use 155 mm square mesh cod ends, resulting in catches of few yellowtail flounder at sizes < 30 cm (H. Stone pers. comm.). Presumably any winter flounder catches in the yellowtail flounder fishery would be of similar size.

Total discards

During the assessment period, 1982-2010, discards of winter flounder on Georges Bank were higher in the U.S. fisheries prior to 1991 and were higher in the Canadian scallop dredge fishery thereafter (Figure B10). Total discards were much higher during 1982-1991, than thereafter, but total discards slowly increased between 1995 and 2010. Total discards averaged 15% of the total landings during 1982-2010.

Catches

Catches increased during 1964-1972, reaching a peak of 4,608 mt in 1972, but then declined to 2,034 mt in 1976 (Figure B11, Table B3). Catches subsequently increased to 4,290 mt in 1981 then gradually declined to a time series low of 842 mt in 1995. Catches increased to 3,328 mt in 2003 then declined to 1,039 mt in 2007, followed by a slight increase to 2,013 mt in 2009. Total catch in 2010 was 1,544 mt.

Historical catches are likely to have been higher than those observed since 1964 because the U.S. landings alone reached a peak of 4,089 mt in 1945, close to the magnitude of the peak catch during 1964-2010 (4,608 mt), and without the addition of discards, at a time when codend mesh sizes were smaller, and landings from international fleets (Figure B11).

Landings-at-age

Length and age composition data are not collected from the landings or discards of Canadian fleets that fish on Georges Bank, but length and age samples from the U.S. landings were collected by market category and quarter during 1982-2010. Samples are collected for eight market categories (Lemon Sole = 1201, Extra Large = 1204, Large = 1202, Large /Mixed = 1205, Medium = 1206, Small = 1203, Peewee = 1207, and Unclassified = 1200). However, the data were binned as Lemon Sole (1201 and 1204), Large (1202 and 1205) and Small (1203, 1206 and 1207) because the three market categories comprised a majority of the landings during 1982-2010. The annual sampling intensity of lengths ranged between 15 mt and 271 mt landed per 100 lengths measured during 1982-2010 (Table B7). Sampling intensity was lowest during 1996-2000. During 1998 and 1999 there were no Lemon Sole samples (the largest market category size) and only one large sample collected during each of these two years (Table B8) although this market category represented 42% and 45% of the total landings, respectively, during this period (Table B9). After 2000, sampling intensity improved substantially and has been highest since 2004 (Table B7, Figure B12). During 1982-2002, landings were dominated by the Large and Small market categories, but during 2002-2008, the landings were dominated by larger fish (Lemon Sole and Large, Table B9), which was reflected in the increased sampling intensity of these larger fish (Figure B9). Landings of Small fish increased after 2006, as the 2006 year class moved through the fishery, and constituted the predominant market category during 2009-2010 (Figure B13).

During most years, biological sampling of the landings was adequate to construct the landings-at-age (LAA) matrix by applying commercial age-length keys to commercial numbers at length on either a quarterly or half-year basis by market category group (Table B10). The LAA matrix was based on that provided in Brown *et al*. (2000), for 1982-1993, and as provided in the 2008 GARM Report (NEFSC 2008) for 1994-2006. LAA data were updated for 2007-2010 using the allocation scheme presented in Table B11. The LAA matrix (nos. in thous.) includes U.S. and Canadian landings during 1982-2010 for fish of ages 1-7+ (Table B11). The U.S. unclassified market category samples and the Canadian landings were assumed to have the same age compositions as the sampled U.S. landings and the U.S. LAA was adjusted by a raising factor to incorporate the Canadian landings.

Large year classes were trackable in the landings-at-age matrix. For example, large numbers of fish from the 1994 cohort were landed as age 1 fish in 1995, as age 2 in 1996 and as age 3 fish in 1997. Landings of age 1 fish were insignificant during most years (Table B11). During 1982-1984, the landings were dominated by age 3-5 fish and were dominated by age 2-4 fish during 1985-2000. Since 2001, the landings have returned to a predominance of age 3-5 fish. In part, this change was due to a codend mesh size increase (to 6.5 in. square or diamond mesh) occurred in the Georges Bank bottom trawl fishery for groundfish in August of 2002.

Discards-at-age

The annual numbers of lengths sampled by fishery observers, from winter flounder discards in the U.S. bottom trawl and scallop dredge fisheries, were inadequate to characterize discard length compositions during 1989-2000 and 1989-2003 (with the exception of 1997), respectively (Table B12). In addition, length and age composition data for winter flounder discards in the Canadian fisheries are not collected. As a result, U.S. bottom trawl discards-at-age were characterized based on the assumption that fish smaller than the U.S. minimum regulatory size limits were discarded. The minimum size limit for winter flounder in the U.S. bottom trawl fishery was 28 cm during 1986-April, 1994 and has been 30 cm since then. Examination of survey length-at-age data indicates that fish of this size are one year old in the NEFSC fall surveys and two years old in the spring surveys. Therefore, discards-at-age for the U. S. bottom trawl fleet, during 1982-2001, was estimated by dividing the estimated weight of discarded winter flounder from the bottom trawl fleet, during January-June, by the annual mean weights of age 2 fish from the NEFSC spring surveys. Likewise, winter flounder discard weights for July-December were divided by the annual mean weights of age 1 fish from the NEFSC fall surveys. Discards-at-age for the U.S. bottom trawl fleet, during 2002-2010, were estimated by using the discard numbers at length from the NEFOP Database, binned as January-June and July-December, to characterize the proportion discarded at length and ages were determined by applying the NEFSC spring and fall survey age-length keys and length-weight relationships, respectively. Length compositions of discarded fish in the U.S. bottom trawl fishery indicate that for most years during 2002-2010, discarding of all sizes of winter flounder occurred (Figure B14), particularly when Georges Bank winter flounder trip limits were in place during May, 2006 - July 6 of 2009 (5,000 lbs per trip). As of October of 2010, all NE multispecies permit holders that fish on a sector trip were prohibited from discarding legal-sized fish (must land all winter flounder > 30 cm TL).

Length samples of winter flounder discarded in the U.S. scallop dredge fishery were inadequate to characterize discard length compositions during 1989-2003, with the exception of 1997 (Table B12). The post-2003 discard length composition data suggested that, in general, all sizes of winter flounder were discarded in the U.S. scallop dredge fishery, but that catches of winter flounder smaller 30 cm are very low (Figure B15). Similar scallop dredges are used by the Canadian scallop fleet (H. Stone, pers. comm.). The Canadian scallop dredge fleet has been prohibited from landing groundfish since 1996 and winter flounder is a low-value species in CA in relation to cod, haddock and yellowtail flounder (there is no existing directed fishery for winter flounder). Given these considerations, discards-at-age for the both the U.S. and Canadian scallop dredge fisheries were estimated by scaling up the LAA by the ratio of total scallop dredge discards to total landings. During years when sufficient numbers of length samples of winter flounder discards were available, 1997 and 2004-2010 (the 2009 and 2010 discard length samples were combined to derive the 2010 discard length composition), the annual discard length frequency distributions were used to characterize the proportion of discards-at-length for both the U.S. and Canadian scallop dredge fleets and the NEFSC fall survey age-length keys and length-weight relationships were applied to the combined annual discard weights (U.S. and CA) because most of the U.S. discards occurred during the second half of the year.

Discards-at-age (numbers in thous.) were computed for ages 1-7+. Discards occurred across all age categories because they are primarily driven by discarding in the U.S. and Canadian scallop dredge fleets. Numbers of discarded fish shifted from primarily age 2-4 fish during 1982-1997 to age 3-5 fish during 1998-2003 (Table B13). The total numbers of fish discarded were consistently much lower during 2004-2010, when the fishing in Closed Areas I and II was mostly prohibited for groundfish trawlers and limited for scallop fishing. However, the range of ages that were discarded broadened to include mostly ages 2-5. Discards of age 1 fish, which occur primarily in bottom trawl rather than scallop dredge fisheries, were highest during 1982-1985; a time when there was no minimum landings size limit in effect and the minimum codend mesh size was smallest (5.5 inches) for groundfish trawlers. During 1982-2010, the numbers of age 1 discards decreased, presumably because the minimum codend mesh size required in groundfish bottom trawls was increased to 6.5 inches.

Catch-at-age

Components of the catch-at-age (CAA) consisted of the combined U.S. and Canadian landings-at-age and discards-at-age for the U.S. large-mesh and small-mesh bottom trawl fleets plus the U.S. and Canadian scallop dredge fleets, during 1982-2010, for ages 1-7+ (Table B14). During 1982-1984, the CAA contained a broad range of ages, but was dominated by ages 2-5 and had the highest numbers of fish aged 6 and older (Table B15, Figure B16). The CAA changed from this more stable age composition to one dominated by ages 2-4, during 1985-1996. During 2000-2005, the catch composition changed back to a predominance of age 3-5 fish and contained more older fish (ages 6 and older), but at the higher levels observed during 1980-1984. The catches were dominated by age 2-4 fish during 2008-2010 as the 2006 year class was harvested by the fishery (Table B15, Figure B16).

Mean weights-at-age in the catch remained relatively stable during 1985-1996 across most ages, but then declined to a lower level during 1997-2001, for ages 3-5 possibly due, in part, to poor sampling of large fish during part of this time period (Figure B17, Table B16). Mean weights-at-age reached their highest levels during 2003-2007, but then declined through 2010 to some of their lowest levels since 1982.

Term of Reference 2: ***Present survey data being considered and/or used in the assessment (e.g., regional indices of abundance, recruitment, state and other surveys, age-length data, etc.). Characterize uncertainty in these sources of data.***

Research Survey Data

Stratified, random bottom trawl surveys conducted by the NEFSC during the spring and fall provide long time series of fishery-independent indices for Georges Bank winter flounder. The fall and spring surveys have been conducted since 1963 and 1968, respectively, and sampling on Georges Bank has generally occurred during October and April, respectively. The strata set used to calculate abundance and biomass indices from the two NEFSC surveys included offshore strata 13-23 (Figure B18). Stratum 23 was included in the strata set for the 2008 stock assessment because age analyses indicated that most fish within the stratum exhibited the faster, Georges Bank growth type rather than the slower growth type of the other two stocks (NEFSC 2008). Winter flounder catches during NEFSC surveys are also highest in the eastern, Georges Bank portion of stratum 23 (NEFSC 2008). A portion of stratum 23 lies within SA 521, for which commercial catches are assigned to the SNE/MA winter flounder stock. Based on a GIS analysis, 46% of stratum 23 is located within Georges Bank SA 522 and the remainder is located in SA 521. However, more than half (53%) of stratum 23 lies within Closed Area 1 and cannot be routinely fished by trawlers targeting winter flounder. Of the open area portion of stratum 23 which can be fished, 74% lies within SA 521, but this is only a small portion of the total area of SA 521.

The SDWG discussed whether the overlap of stratum 23 with SNE-MA SA 521 was a concern with respect to its effect on biological sampling of commercial catches or assessment model tuning indices. However, because of the differences in growth rates between the two stocks, biological samples from catches in SA 521 are readily assigned to the correct stock, eliminating such concerns. Winter flounder catches during both the spring and fall surveys are very low in the open portion of stratum 23 that lies within SA 521, suggesting that commercial catches of winter flounder from this portion of SA 521 are also likely low, and therefore, are not expected to influence the assessment results. As a long-term solution to this issue, splitting stratum 23 into two strata, is planned for the 2011 fall survey.

Relative abundance and biomass indices of Georges Bank winter flounder derived from Canadian stratified random bottom trawl surveys, conducted in strata 5Z1-4 (Figure B19) during February by the Maritimes Region staff from the Division of Fisheries and Oceans, were also included in the assessment. The survey design and sampling protocols are provided in (Chadwick et al. 2007).

Beginning in 2009, the NEFSC SRV *Albatross IV* was replaced with the SRV *Henry B. Bigelow*. The new vessel is quieter and the reduced spacing between the rockhoppers on the footrope has improved the catchability of winter flounder. In order to extend the NEFSC spring and fall survey time series beyond 2008, stock-specific, length-based vessel calibration factors were applied to the Bigelow catches of Georges Bank winter flounder to convert them to Albatross equivalents. The data and methods used to estimate the calibration factors are described in Appendix B3. The aggregate catch number calibration factor for winter flounder, for combined seasons, is 2.490 and the aggregate catch weight factor, for combined seasons, is 2.086 (Miller et al. 2010). A fourth order polynomial model fit to data for the Georges Bank stock region, incorporating a mean ratio of the vessel swept areas of 0.5868 (Bigelow to Albatross), was used to calculate the calibration factors-at-length (Figure B20) that were used to convert the 2009-2010 Bigelow survey indices to Albatross units for use in the VPA model (Table B17).

Relative biomass (stratified mean kg per tow) and abundance (stratified mean number per tow) indices are presented for the NEFSC spring (April, 1968-2010) and fall (October, 1963-2010) bottom trawl surveys , as well the Canadian spring bottom trawl surveys (February, 1987-2010, (Table B18). NEFSC survey indices prior to 1985 were standardized for gear changes (weight = 1.86 and numbers = 2.02, Sissenwine and Bowman 1978) and trawl door changes (weight = 1.39 and numbers = 1.4, Byrne and Forrester 1991).

Despite considerable inter-annual variability, the NEFSC fall survey relative abundance indices showed an increasing trend during the 1970's, followed by a declining trend during the 1980s to a time series low in 1991 (Figure B21). Thereafter, relative abundance increased through 2001 then declined to a level below the 1963-2009 median during 2005-2007. In 2009, fall relative abundance reached the second highest point in the time series, but declined drastically in 2010 to a level slightly below the time series median. Trends in the NEFSC spring survey relative abundance indices exhibited more inter-annual variability, but were similar to the fall survey time series after 1982. NEFSC spring survey abundance indices were at record low levels during 2004-2007. The second highest abundance index of the time series occurred in 2008. However, most of the fish were caught at two consecutively sampled stations and relative abundance was much lower in 2009 and was at the time series median level in 2010. Relative abundance trends in the Canadian survey were similar to those in the NEFSC spring survey during most years but were of greater magnitude during blocks of years (1988-1990 and 1993-1997). Similar to relative abundance indices from the NEFSC spring surveys, indices from the Canadian surveys were at the lowest levels observed during 2005-2007 but were well below the time series median in 2009 and 2010.

In order to estimate catchability coefficients for each survey (q) in the VPA, minimum population size estimates were computed based on swept areas of 0.011 nmi^2, for NEFSC surveys conducted by the Albatross and Delaware, and 0.012 nmi^2 for the CA surveys. During NEFSC and CA surveys, tows are conducted for 30 minutes, between winch lock and re-engage, at a target speed of 3.5 knots (Azarovitz 1981; Chadwick et al 2007). Minimum population sizes-at-age (000's) included in the VPA included: the U.S. fall (1981-2010, ages 0-6 lagged forward one year and age, Table B19) and spring bottom trawl surveys (1982-2010, Table B20) and the Canadian spring bottom trawl surveys (1987-2010,Table B21). Age samples of winter flounder are not collected during Canadian bottom trawl surveys so the NEFSC spring survey age-length keys, augmented during some years with commercial age-length keys from the first quarter of the corresponding year (when larger fisher were caught), were used to partition stratified mean numbers-at-length from the Canadian surveys into numbers-at-age. Although the numbers-at-age were highly variable, large cohorts appeared to track through the numbers-at-age matrices, for the NEFSC surveys, for the 1980, 1987, 1994, 1998-2001, and 2006 cohorts (Figure B22). Age truncation occurred between 1983 and 1997, during which time the population was dominated by four age groups rather than seven or more. During 1997-2004, the age structure improved but has since become truncated again. Both the U.S. and Canadian spring surveys showed reduced numbers of age 1-3 fish (and age 4 fish in the CA surveys) during 2000-2007. The Canadian spring survey did not show the same magnitudinal increase in age 1-6 fish that was evident in the NEFSC spring surveys during 2008-2010.

Term of Reference 3: ***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 area-swept biomass estimates. Investigate if implied survey gear or catchability estimates are reasonable. Include a historical retrospective analysis to allow a comparison with previous assessment results.***

Model input data

A series of VPA model runs was conducted for the current assessment using the NOAA Fisheries Toolbox (NFT) ADAPT VPA (Gavaris 1988) version 3.0.3. (NFT 2010) and data for 1982-2010. Retrospective analyses, for terminal years 2001-2009, were conducted for each model run. Input data and descriptions of the different model formulations are presented in Table B22. An initial population analysis was conducted to provide a "bridge" from the 2008 GARM assessment results (NEFSC 2008) by updating the 2008 model configuration. However, the model results indicated that stock estimates for age 2 fish were no longer estimable in the terminal year +1 (CV = 1). As a result, all subsequent model runs included stock estimates for ages 3-6 in 2011. Run 4 included the new three-year moving window maturity schedule (described above in the Growth and Maturity section) and the addition of Canadian scallop dredge discards with an M assumed of 0.2. The Run 5 model formulation was the same as for Run4, but included the Working Group's recommended increase in M to 0.3 (see M section above).

Sensitivity Run 1 evaluated the effect of the new maturity schedule on the SSB estimates. Sensitivity Runs 2 and 3 evaluated the effects of omitting Canadian spring survey as a tuning index and down-weighting of the Canadian survey residuals, respectively. The Canadian survey was responsible for the highest percentage (43%) of the total variance of all three tuning indices. In particular, Canadian survey indices for ages 1-3 comprised the highest percentages of the total variance of indices from all three surveys (7.1%, 8.3% and 9.8%, respectively). However, the Working Group recommended against selecting subsets of ages from the tuning indices. For Sensitivity Run 3, the Canadian survey residuals were down-weighted by 0.42, which was computed as the squared average of the ratios of the mean CVs of each of the U.S. survey indices to the average CV of the Canadian survey indices.

Results

There was little difference in the VPA estimates of average F, SSB and age 1 recruitment for the updated 2008 GARM run versus Run 4 (Figure B23). The latter run included the the new maturity schedule and the addition of discards from the CA scallop dredge fleet. The largest difference in F, SSB and R estimates between these two runs and Run 5 was attributable to the increase in M from 0.2 to 0.3 . For Run 5, estimates of F were lower and SSB and R estimates were higher than for the other two runs (Figure B23). The result of applying the new maturity schedule was a 4.5-28.4% reduction, 14% on average, in the annual SSB estimates during 1982-2010 (Figure B24).

Model diagnostics

Trends in the residuals patterns were evident for a number of ages within each of the three sets of VPA calibration indices, with variability by age and year. For example, residuals trends from NEFSC spring surveys were the worst for age 2 and age 3 fish. Residuals were positive for age 2 fish during 1990-1996, and for age 3 fish during 1983-1987, but were negative for age 3 fish during 2001-2007 (Figure B25). The Canadian spring survey indices for ages 2-4 showed major residuals trends (Figure B26), both positive and negative, but the patterns differed from those evident in the NEFSC spring surveys. For example, age 3 and age 4 fish showed similar residuals trends; positive during 1988-1991 and 1993-1997, but negative during 1998-2010. Residuals trends for the NEFSC fall survey abundance indices were the worst for older fish,ages 5-7 (actually ages 4-6 lagged forward one year and age) and were generally positive from 2002 or 2003 onward (Figure B27).

VPA estimates of survey catchability coefficients (q), by age, indicated that catchabilities for all three surveys generally increased with age (Figure B28). Catchabilities were higher for the NEFSC fall surveys than the NEFSC spring surveys (e.g., q = 0.20 and 0.28 for age 6, respectively), but qs-at-age between the two surveys were not significantly different. Catchabilities for the Canadian spring surveys can be compared across ages but not between surveys because the vessels and gear were different. The catchabilities of ages 1-3 fish were significantly lower than for ages 5-7+ fish (Figure B28).

Retrospective analyses

Similar to the 2008 GARM assessment results, very mild retrospective patterns were present for terminal year estimates of fishing mortality (overestimation during 2006-2009 and underestimation during 2002-2005) and spawning stock biomass (underestimation during 2006-2009 and overestimation during 2002-

2005, Figure B29). There was no retrospective pattern for terminal year age 1 recruitment, but the estimates were highly variable (Figure B29).

Rrelative differences in the estimates of average F, SSB and age 1 recruitment, during year *t* (for 2001-2009) versus 2010, are presented in Figure B30. Run 5 was selected as the final model run because the range of retrospective errors in F and SSB was narrower than for Sensitivity Runs 2 and 3 (Table B23). For Run 5, the retrospective error in fishing mortality ranged from -48% in 2002 to +42% in 2009 and retrospective error in SSB ranged from -13% in 2008 to +43% in 2002 (FigureB30).

Estimates of fishing mortality, spawning stock biomass and recruitment

Estimates of January 1 population size (numbers, 000's), average fishing mortality rates (F on ages 4-6), and spawning stock biomass (mt), from the final VPA model run, are presented in Tables B24-B26, respectively. Fishing mortality rates were highest during 1984-1993, ranging between 0.57 and 1.17, then declined to levels ranging between 0.31 and 0.51 during 1994-1998 (Figure B31, Table B27). Fishing mortality rates were low (0.26-0.27) during 1999 and 2000, then increased rapidly to 0.85 in 2003 and was followed by a rapid decline to the second lowest level in the time series (0.20) in 2006. Fishing mortality increased slightly during 2007-2009, but then declined to 0.15 in 2010.

SSB declined rapidly from a time series peak of 17,380 mt in 1982 to 6,256 mt in 1985, then increased slightly through 1987 to 8,082 mt (Figure B31, Table B27). After 1987, SSB declined again to a time series low of 3,424 mt in 1995. SSB subsequently increased to 13,790 mt in 2000, but then declined to 5,305 mt in 2005. Thereafter, SSB increased and totaled 9,703 mt in 2010.

Trends in age 1 recruitment showed several periods of rise-and-fall. Recruitment increased from 8.3 million fish in 1983 to a time series peak of 26.3 million fish in 1988, and then declined to 5.2 million fish in 1993 (Figure B31, Table B27). Recruitment increased again to fairly high levels during 1995-1999 (16.2-22.8 million fish) then declined to the second lowest level on record (5.5 million fish) in 2004. Recruitment increased to 18.8 million fish in 2008, but then declined to the lowest level in 2009 (4.0 million fish). Recruitment increased to a very high level (22.5 million fish) in 2010; an estimate that was based on the partial recruitment value for age 1 fish multiplied by the fully-recruited F. The 2011 recruitment value is uncertain because it represents the geometric mean of the 2003-2009 recruitment values. Bootstrapped estimates of the 2011 stock sizes-at-age and the 2010 fishing mortality rates-at-age are presented in Tables B28 and B29, respectively.

A comparison of the estimates of F, SSB and R, from the final model (Run 5) versus the two sensitivity runs indicated, that in recent years, slightly higher F and lower SSB and R values were estimated when the CA spring survey was included as a tuning index (in the final model, Run 5) rather than when the CA survey was omitted or downweighted (Figure B32). As discussed previously in the Retrospective Analyses section, the two sensitivity runs resulted in increases in retrospective error in F and SSB in comparison to Run 5.

TOR 4: ***Perform a sensitivity analysis which examines the impact of allocation of catch to stock areas on model performance (in TOR-3).***

The SDWG interpretation of TOR4 was that the variance of the commercial landings due to the 1995 and later area-allocation scheme should be used as the basis for estimating the magnitude of landings that might be lost or gained for the stock-specific assessments, and that the assessment models should be run with those potential biases incorporated and the results presented. For the Georges Bank stock, annual catches consisted of the U.S. landings and discards and the Canadian landings and discards. Precision estimates for the Canadian landings and discard estimates were not provided by the CA DFO, so they were assumed to be the same as the precision estimates for the US landings and discards.

For the Georges Bank winter flounder stock, total landings for 1995-2010 have a calculated Proportional Standard Error (PSE; due to the aforementioned commercial landings area-allocation procedure) ranging from 0.7% to 1.3% (Table B30). The 1995-2010 mean PSE of 0.9% was substituted for the 1982-1994 PSEs of the landings. The total discard PSEs during 1995-2010 ranged from 1% to 56%. The 1995-2010 mean PSE of 26% was substituted for the PSEs of the 1982-1994 discards. Because the PSEs for the landings are low, and the landings accounted for 69-94% of the total catch during 1982-2010, the total catch-weighted annual PSEs ranged from 1.2% to 8.2% and averaged 3.9% (unweighted) for the 1982-2010 time series.

The SDWG developed an exercise using the 2008 GARM-III assessment data and ADAPT VPA model in an initial response to TOR4 and concluded that the application of a annually varying unidirectional "bias-correction" in such an exercise provides stock size estimates and BRPs that scale up or down by about the same average magnitude as the gain or loss (SDWG52 WP3).

Since development of SDWG WP3, the SDWG concluded that the calculated variance of the area-allocated commercial landings likely underestimates the true error. More work was done to estimate the error in the commercial landings due to misreporting of commercial landings to statistical area at allocation level “A” reporting level in mandatory Vessel Trip Reports (Palmer and Wigley MS 2011). Vessel monitoring system (VMS) positional data from northeast United States fisheries for 2004-2008 were used to validate the statistical area fished and stock allocation of commercial landings derived from the VTRs. The accuracy of the VMS method relative to the VTRs was assessed using haul locations and catch data recorded by at-sea observers. This work was performed for several New England groundfish species. The perceived under-reporting of statistical areas in the VTR data led to minor (< 5%) differences in the overall species landings allocations. Only nine stocks in the five year time-series exhibited differences in stock allocations exceeding 2.0% (2004: northern and southern silver hake, ± 3.0%; 2006: northern and southern windowpane flounder, ± 4.7%; 2007: Georges Bank winter flounder, 2.4%; 2008: Georges Bank winter flounder, 2.4%, Southern New England/Mid-Atlantic winter flounder, -3.2%, and northern and southern windowpane flounder, ± 3.4%). Given the magnitude of these errors, the SDWG elected to update the exercise by adding an additional 5% PSE to the PSE values shown in Table B30 for the Georges Bank total landings during 1995-2010. This increased the 1995-2010 average landings PSE from 0.9% to 5.7%, and increased the average 1982-2010 catch PSE from 4.0% to 6.2%, with a range of 2.7% in 1983 to 13.7% in 2010.

The catch in the final assessment model was increased/decreased by the annually varying catch PSEs and models were re-run to provide an additional measure of the uncertainty in assessment estimates. As noted in SDWG WP3, the application of a annually varying "bias-correction" in one direction in such an exercise provides stock size estimates that scale up or down by about the same average magnitude as the gain or loss. For the final VPA model results, fishing mortality did not change, on average (out to three decimal places), and the range in 2010 F was 0.154 to 0.162. SSB changed by – 1.0% and +7.9%, on average, and the range in 2010 SSB was 9,636 mt - 10,504 mt (Figure B33).

Term of Reference 5: ***Examine the effects of incorporating environmental factors in models of population dynamics (e.g., spring water temperatures in an environmentally-explicit stock recruitment function).***

Winter flounder spawn in winter and early spring in estuaries along the mid-Atlantic, southern New England and Gulf of Maine (Able and Fahay 2010) as well as in continental shelf waters on Georges Bank during March-May (Smith 1985). There is also recent evidence of more coastal spawning in both Southern New England (Wuenschel et al. 2009) and the Gulf of Maine (Fairchild et al. 2010). In southern New England, Manderson (2008) found that overall recruitment was linked to spring temperatures, presumably by acting on larvae, settlement stage, and/or early juveniles. Further, Manderson (2008) found that young-of-the-year abundance among 19 coastal nurseries became more synchronized in the early 1990's and argued that increased frequency of warm springs was creating coherence in early life stage dynamics among local populations.

The specific mechanism linking temperature to recruitment was not defined by Manderson (2008), but temperature is an important parameter in many ecological processes affecting winter flounder. In a mesocosm study, Keller and Klein-MacPhee (2000) found that winter flounder egg survival, percent hatch, time to hatch, and initial size were significantly greater in cool mesocosms. Further, mortality rates were lower in cool mesocosms and related to the abundance of active predators. In the laboratory, Taylor and Collie (2003) found that consumption rates of sand shrimp were lower at lower temperatures implying lower predation pressure at colder temperatures. In the field, Stoner et al. (2001) found that settlement stage winter flounder prefer colder waters and that the importance of temperature in defining juvenile habitat decreases through ontogeny. Thus, temperature has multiple effects on the early life history of winter flounder and colder temperatures in general lead to higher survival and recruitment.

The relationship between winter flounder recruitment and temperature identified by Manderson (2008) did not include the effect of population size. The relationship between stock size and subsequent recruitment is generally poor in marine fishes (Rothschild 1986) but can have explanatory power. To examine the combined effect of environment and spawning stock biomass on recruitment, the goal was to develop environmentally-explicit stock-recruitment relationships that include temperature and related environmental variables for the three stocks of winter flounder. As a basic framework, the approach of Hare et al. (2010) was followed. The resulting models could be used in short-term forecasts based on fishing and temperature scenarios (fixed patterns of temperature variability over several years) and long-term forecasts based on fishing and temperature projections from general circulation models. The methods and results of the analysis are described in Appendix B2.

The conclusion from the analysis was that recruitment in the coastal stocks of winter flounder (GOM and SNE-MA) were linked to air temperatures during winter, when spawning occurs, but there was no evidence for an air temperature effect on recruitment in the Georges Bank stock; the environmentally-explicit models (which also included a Gulf Stream index) did not provide a better fit compared to the standard stock recruitment model. The Georges Bank stock experiences water temperatures that are affected by both local air temperatures and more importantly, large-scale advective supplies of relative cold, fresh water associated with the Labrador Current. Examining other environmental variables which may affect recruitment in the Georges Bank stock (e.g., hydrographic circulation patterns on Georges Bank in relation to larval abundance) is listed below as a future research recommendation.

Term of Reference 6: ***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}, $B_{THRESHOLD}$, and F_{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.

Existing biological reference points (BRPs)

The specification of FMSY and BMSY reference points relies on a stock-recruitment relationship. As a result, the 2008 GARM Biological Reference Point Review Panel (NEFSC 2008) concluded that MSY-based BRPs should be adopted when the stock-recruitment relationship is informative, and if not, then the Panel recommended the use of F40%MSP as a proxy for FMSY, similar to the previous recommendation from a separate BRP Working Group for many of the groundfish stocks (NEFSC 2002a), and a BMSY proxy computed using the non-parametric, empirical approach.

For Georges Bank winter flounder, the 2008 GARM BRP Review Panel (O'Boyle et al. 2008) concluded that the Beverton-Holt stock-recruit relationship, derived using data from the VPA model, was uninformative regardless of whether the model was fit without a prior (h=1) or with a prior (the fit was highly dependent on the assumed prior on unfished recruitment, R_0). Thus, the Panel recommended that a non-parametric empirical approach be used to estimate biological reference points based on: 1) the final VPA model results, 2) the estimate of F40%MSP as a proxy for F_{MSY} (derived from a per-recruit model using the most recent five-year average of fishery selectivity and weights-at-age and the maturity-at-age time series average), and 3) a long-term (100-year) stochastic projection using the cumulative distribution function of observed recruitment (1983-2007 recruitment at age 1, the 1982-2006 year classes) to estimate MSY and SSBMSY40%. The existing BRPs, F40% and SSB40%, were adopted at the 2008 GARM (NEFSC 2008) and were promulgated in 2009 in Amendment 16 to the Northeast Multispecies Fishery Management Plan (NEFMC 2009). The existing biomass target is SSBMSY at 40% MSP (= 16,000 mt) and the minimum biomass threshold is 50% of the target (= 8,000 mt). The fishing mortality threshold is F40%MSP (= 0.26). Amendment 16 defines the fishing mortality target as the mortality associated with the Annual Catch Limit (ACL).

Candidate biological reference points

For Georges Bank winter flounder, two sets of candidate BRPs (i.e., FMSY and SSBMSY versus F40% and SSB40%) were brought forward from the current assessment for review by the SARC because the SDWG could not reach consensus on whether the stock-recruit relationship from the Beverton-Holt model was informative, and consequently, whether FMSY was well-estimated. Both sets of BRPs were estimated similar to the methods used for the 2008 GARM (NEFSC 2008), as summarized in the preceeding paragraph.

FMSY was estimated from a Beverton-Holt model which incorporated R (age 1) and SSB estimates from the final VPA model (1982-2009 year classes) with an assumed prior on steepness (h = 0.8 and SE = 0.09, based on the values reported for Pleuronectids in Myers *et al*. (1999)). . In addition, a per-recruit model (Thompson and Bell 1934) was used to estimate an F_{MSY} proxy of $F_{40\%\ MSP}$. Input data to both models included the most recent five-year averages (2006-2010) of fishery selectivity-at-age, proportion mature-at age, and weights-at-age from the final VPA model (Table B31).

Parameter estimates from the Beverton-Holt model are shown in Table B32. Similar to the 2008 GARM results, the steepness parameter for the Beverton-Holt model could not be estimated (h=1) without assuming a prior. This constant recruitment even at low spawning stock sizes is not theoretically feasible. When the steepness prior was set to 0.8, with a standard error of 0.09, the h estimate was 0.85 (CV = 0.08; 80% CI = 0.74, 0.94) and the FMSY estimate was 0.50 (CV=0.22; 80% CI = 0.39, 0.69). Precision estimates were obtained from an MCMC analysis with 1,000 realizations (100,000 MCMC iterations with a thinning rate of 100). The steepness log-likelihood profile indicated that the steepness prior was highly influential in determining the FMSY estimate (Table B33). Both sets of candidate BRPS presented to the SARC are shown in Table B34, along with the existing BRPs.

The SARC expressed concerns about how well the Myers *et al*. (1999) steepness value for Pleuronectids was estimated and that the values of M upon which their models were based were lower (≤ 0.2) than the value of 0.3 used in the SARC 52 winter flounder assessments. The SARC noted that the stock-recruitment data for the Georges Bank stock was less informative than the SNE/MA data for predicting recruitment at low spawner levels, making direct estimation of the spawner-recruit relationship difficult without external information. The SARC also concluded that steepness values should be similar between winter flounder stocks. Therefore, the steepness log-likelihood profiles of the two stocks (Table B33 for the Georges Bank stock) were used in selecting fixed values for steepness with which to estimate FMSY for each stock. Fixed values of steepness were chosen that were as similar as possible between the stocks, but which also provided good fits to the stock-recruit data for each stock. Steepness values that are within two units of the minimum AIC were considered to be realistic values for each stock (Burnham and Anderson, 2002). Therefore, the SARC recommended that steepness be set at the largest value such that ΔAIC = 2, for the SNE/MA stock (steepness fixed at 0.61, Figure B34), and at the smallest value such that ΔAIC = 2 for the Georges Bank stock (steepness fixed at 0.78, Figure B34). The final candidate FMSY estimate resulting from fixing steepness at 0.78 is 0.42 (Table B33). Precision estimates for FMSY were not possible due to fixing the steepness parameter. Results from the model fit and standardized residuals are shown in Figure B35. Trends in the residuals alternate between positive and negative for most of the time series. Estimates of SSBMSY and MSY, and their associated precision, were estimated using the method described above for the 2008 GARM; a 100-year stochastic projection that incorporated the parameter estimates from the Beverton-Holt model and the cumulative distribution function of observed recruitment (1983-2010 recruitment at age 1, the 1982-2009 year classes). Candidate BRPs estimated for the Georges Bank winter flounder stock which were used to determine 2010 stock status were: FMSY (Fthreshold) = 0.42; SSBMSY (Btarget) = 11,800 mt; ½ SSBMSY (Bthreshold) = 5,900 mt and MSY = 4,400 mt (Table B35).

Term of Reference 7: *Evaluate stock status (overfished and overfishing) with respect to the "new" BRPs (from TOR 6), and with respect to the existing BRPs (from a previous accepted peer review) whose values have been updated.*

Stock status

In 2010, overfishing was not occurring because the 2010 fishing mortality rate (= 0.15) was below the value of FMSY (= 0.42, Table B35). The stock was also not overfished in 2010 because spawning stock biomass in 2010 (= 9,703 mt) was above the SSB_{MSY} threshold (= 5,900 mt, Table B35, Figure B36).

The results of a bootstrap analysis (1,000 iterations) suggested that the 2010 estimates of average F (on fully recruited ages 4-6) and spawning stock biomass were fairly precise with CVs of 20% and 24%, respectively. There was an 80% probability that the 2010 F estimate was between 0.12 and 0.21 and that the 2010 SSB estimate was between 7,304 mt and 12,578 mt (Figure B37).

In the current assessment, the assumed value for M was increased from 0.2 to 0.3. As a result, the SDWG concluded that a comparison of the 2010 F and SSB estimates from the current assessment with the existing reference points was not appropriate.

The revised assessment model alters the historical perception of stock status. Four changes from the previous assessment are: 1) a change of M from 0.2 to 0.3 and 2) a new maturity schedule, 3) the addition of Canadian discards, and 4) a change to MSY-based BRPs rather than proxies. Based on the results from the revised assessment model, the stock was overfished during 2004 aqnd 2005. During 2006-2010, spawning stock biomass was above the new biomass threshold of 5,900 mt, but did not reach the new biomass target of 11,800 mt. This contrasts with the 2008 assessment which indicated the stock was overfished in 2007.

Term of Reference 8: ***Develop and apply analytical approaches and data that can be used for conducting single and multi-year stock projections and for computing candidate ABCs (Acceptable Biological Catch; see Appendix to the TORs) under a set of alternative harvest scenarios. If the stock needs to be rebuilt, take that into account in these projections.***

- ***a. Provide numerical short-term projections (3-5 yrs, or through the end of the rebuilding period, as appropriate). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. In carrying out projections, consider a range of assumptions about the most important uncertainties in the assessment (e.g., terminal year abundance, variability in recruitment). Take into consideration uncertainties in the assessment and the species biology to describe this stock's vulnerability (see "Appendix to the SAW TORs") to becoming or remaining overfished, and how this could affect the choice of ABC.***
- ***b. Develop plausible hypotheses (e.g., mixing among the three stocks) which might explain any conflicting trends in the data and undertake scenario analyses to evaluate the consequences of these alternate hypotheses on ABC determination.***

Projections

Stochastic medium-term projections of future stock status, during 2011-2017, were conducted based on results from the final VPA model run and the candidate BRPs using AGEPRO software (v. 3.3) from the NOAA Fisheries Toolbox (NOAA 2009). Maturity-at-age and mean weights and fishery selectivity patterns-at-age, estimated for the most recent 5 years of the assessment (2006-2010), were included in the projections to reflect current conditions in the stock and fishery (Table B31). The projections assumed that a catch of 2,118 mt (for the FMP Framework 44 fishing year beginning May 1) would be landed as the calendar year catch in 2011. The projections incorporated uncertainty in the current population estimate, via bootstrap replicates (N=1,000), and variability in predicted recruitment. A parametric Beverton-Holt model with log-normal error was used and recruitment variability was generated by randomly sampling from the estimated error distribution of the fitted stock–recruitment model.

The regulations require rebuilding of the Georges Bank stock, with at least 75% probability, by 2017. The projections indicated that rebuilding to SSBMSY (= 11,800 mt) is expected to be achieved with 78% probability in 2012 and 93% probability in 2012 when fishing at 75% of FMSY (=0.315) with a catch of 2,118 mt in 2011 (Figure B38). Projected SSB, during 2011-2017, and catches, during 2012-2017, and their 10% and 90% confidence intervals are shown in Figure B39.

Stock Vulnerability

Appendix to the SAW TORs: "*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).*"

Vulnerability, productivity and susceptibility of the Georges Bank winter flounder stock using several methods. Uncertainty was evaluated using model estimates of precision and qualification of other uncertainties. The age-based VPA model and associated MSY reference point evaluations provide a relatively comprehensive and synthetic evaluation of vulnerability that is entirely consistent with stock status determination and projection. With respect to status determination, vulnerability and susceptibility were accounted for with regards to estimation of F in 2010, but precision estimates for F_{MSY} were not possible due to the use of a fixed steepness value in the Beverton-Holt stock-recruit model. Stock vulnerability and susceptibility were also accounted for in the stock rebuilding projection. All components of productivity (reproduction, individual growth, and survival) were also explicitly accounted for in stock status determination and projections. Reproduction was monitored as age-1 recruitment, and projected as a function of SSB (the product of abundance, weight- and maturity-at- age). Individual growth was monitored as empirical size at age, and projected as recent mean size at age. Survival was accounted for based on model estimates of fishing mortality and selectivity as well as assumed natural mortality, which was informed by tagging analysis.

Uncertainties that were not accounted for by the VPA and reference point models were evaluated using model diagnostics. Standard model diagnostics (e.g., residual analyses, retrospective analyses) were used for model validation. Retrospective patterns were not problematic for Georges Bank winter flounder.

Vulnerabilities that were not accounted for from the assessment and reference point models were evaluated using exploratory modeling, habitat observations and testing the influence of environmental factors on recruitment dynamics. The Georges Bank winter flounder stock is harvested primarily by US bottom trawlers engaged in the large-mesh, multispecies groundfish fisheries. Bycatch and discards are monitored and managed through Annual Catch Limits with Accountability Measures for exceeding those limits. However, a small portion of the stock (5-17% of the total catch during 2004-2010) is not regulated by the US, yet is susceptible to fishing (i.e., incidental catches) by the Canadian scallop dredge and groundfish bottom trawl fleets. Winter flounder discards in the latter fleet are unknown.

An additional consideration of vulnerability and productivity are the implications of increased natural mortality from predation. Consumption of winter flounder by other fishes, birds and marine mammal predators, particularly seals, may be increasing if these predator populations are increasing.

Potential for stock mixing

Historical tagging studies (e.g., Howe and Coates 1975) indicate that there is limited mixing of fish among the three current stock units, with about 1%-3% between the GOM and SNE/MA, about 1% between GBK and SNE/MA, and <1% between GOM and GBK. Historical meristics studies based mainly on fin ray counts also indicate a separate GBK stock (Kendall 1912; Perlmutter 1947) or separate GOM, GBK, and SNE stocks (Lux et al. 1970; Pierce and Howe 1977). Growth and maturity studies also support the distinction of at least three stock areas (Lux 1973; Howe and Coates 1975; Witherell and Burnett 1993), with GBK growing and maturing the fastest and GOM fish the slowest.

The SDWG has initiated research pursuing the use of a more complex model (i.e., Stock Synthesis) to maintain separate fishery and survey catch for the three current stock units, while allowing a small amount (a few percent) of exchange between the stock units based on information from historical tagging. However, development of that research has not progressed sufficiently to be made available for peer review at this time.

Term of Reference 9: ***Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in recent SARC reviewed assessments and review panel reports. Identify new research recommendations.***

Research recommendations from previous assessments

2002 GARM

1. *Investigate whether NEFSC survey stratum 23 includes winter flounder from the Georges Bank stock.*

Most fish in stratum 23 exhibited much faster Georges Bank-type growth rates, so stratum 23 has been included in stock assessments since the 2008 GARM.

2. Request additional observer coverage of GB SD and BT fisheries.

As of 2004, sea day allocations have been based on effort patterns in the scallop dredge and large mesh (codend mesh > 5 in.) bottom trawl fleets and NEFOP funding has increased.

2005 GARM

1. Include discards in future assessments.

US fishery discards were included in the 2008 GARM assessment.

2008 GARM

1. *Explore assessment approaches that consider all three stocks with interaction amongst them.*

An SS3 modeling exercise to explore this approach is currently in progress at the NEFSC (see TOR 8, Potential for stock mixing).

2. Examine why the resource has declined when the harvest has not exceeded MSY (3,500 mt at the 2008 GARM) since 1984.

Total biomass estimates from 2005 assessment (ASPIC model results), indicated that biomass was highest prior to 1982, the initial year of the VPA.

SARC 52 research recommendations

The following research recommendations are listed in order of priority, by topic, in order to focus on research which will provide the most benefit to improving the stock assessment:

Stock-recruitment relationships

Revise the NEFSC assessment software to include the ability to model S-R functions including environmental factors with errors/probabilities.

Further explore the relationship between large scale environmental forcing (e.g., temperature, circulation, climate) for effects on life history, reproduction, and recruitment in the Georges Bank stock.

Explore development of an index of winter flounder larval abundance based on MARMAP, GLOBEC, etc. time series.

Improvements to landings data

Investigate ways to improve compliance to help VTR reporting. Currently about 300 of the 1500 permitted vessels consistently under-report the number of statistical area fished.

Aging

Investigate the feasibility of port samplers collecting otoliths from large and lemon sole instead of scales because of problems under-ageing larger fish.

Reproduction

Investigate the use of periodic gonad histology studies as a check to make ensure maturity estimates are accurate, with particular attention to obtaining sufficient samples from the Georges Bank stock.

Investigate the skipped spawning percentage for each stock, and estimate interannual variation when sufficient data have been collected.

Fishery-independent surveys

Encourage support for industry-based surveys, which can provide valuable information on stock abundance, distribution, and catchability in research surveys that are independent of and supplemental to NMFS efforts.

Modeling

Explore use of a more complex Stock Synthesis model with small rates of migration between stocks.

Consumption

Develop a time series of winter flounder consumption by the major fish predators of winter flounder.

References

Able KW, Fahay MP. 2010. Ecology of Estuarine Fishes: Temperate Waters of the Western North Atlantic. John Hopkins University Press, Baltimore.

Ampela K. 2009. The diet and foraging ecology of gray seals (*Halichoerus grypus*) in United States waters. Ph. D. dissertation. The City University of New York. 176 p.

Anonymous. 2009. Independent Panel review of the NMFS Vessel Calibration analyses for FSV *Henry B. Bigelow* and R/V *Albatross IV*. August 11-14, 2009. Chair's Consensus report. 10 p.

Azarovitz T.R. 1981. A brief historical review of the Woods Hole Laboratory trawl survey time series. Pages 62-67 in W.G. Doubleday and D. Rivard, editors. Bottom trawl surveys. Canadian Special Publication of Fisheries and Aquatic Sciences 58.

Berry RJ, Saila SB, Horton DB. 1965. Growth studies of winter flounder *Pseudopleuronectes americanus* (Walbaum), in Rhode Island. Trans. Am. Fish. Soc. 94: 259-264.

Beverton RJH, Holt SJ. 1957. On the dynamics of exploited fish populations. Chapman and Hall, London. Facsimile reprint. 1993. 533 p.

Bigelow HB, Schroeder WC. 1953. Fishes of the Gulf of Maine. Fishery Bulletin of the Fish and Wildlife Service 74 Volume 53. 577 p.

Brodziak J, O'Brien L. 2005. Do environmental factors affect recruits per spawner anomalies of New England groundfish? ICES J. Mar. Sci. 62:1394-1407.

Brodziak JTK, Overholtz WJ, Rago P. 2001. Does spawning stock affect recruitment of New England groundfish? Can J Fish Aquat Sci. 58: 306-318.

Brooks EN, Miller TJ, Legault CM, O'Brien L, Clark KJ, Gavaris S, Van Eekhaute L. 2010. Determining length-based calibration factors for cod, haddock, and yellowtail flounder. TRAC Ref. Doc. 2010/08.

Brown R. 2009. Design and field data collection to compare the relative catchabilities of multispecies bottom trawl surveys conducted on the NOAA ship *Albatross* IV and the FSV *Henry B. Bigelow*. NEFSC Bottom Trawl Survey Calibration Peer Review Working Paper. Northeast Fisheries Science Center, Woods Hole, MA. 19 p.

Brown, R. W., J. M. Burnett, G. A. Begg, and S. X. Cadrin. 2000. Assessment of the Georges Bank winter flounder stock, 1982-1997. Northeast Fish. Sci. Cent. Ref. Doc. 00-16; 88 p.

Brownie C, Anderson DR, Burnham KP, Robson DS. 1985. Statistical inference from band recovery data. Fish and Wildlife Serv. Resource Publ. 156.

Burns, T. S., R. Schultz, and B. E. Brown. 1983. The commercial catch sampling program in the northeastern United States. Can. Spec. Pub. Fish. Aquat. Sci. 66.

Burnham, K. P., and Anderson, D.R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach, 2nd ed. Springer-Verlag. New York, NY, USA. 488 p..

Byrne, C.J. and J.R.S. Forrester. 1991. Relative fishing power of two types of trawl doors. Northeast Fish. Sci. Center Stock Assessment Workshop (SAW 12). 8 p.

Chadwick EMP, Brodie W, Colburne E, Clark D, Gascon D, Hurlbut T. 2007. Atlantic zone monitoring program bulletin. No. 6 http:// www.meds-sdmm.dfo-mpo.gc.ca/zmp/main_zmp_e.html 68 p.

CSAS [Canadian Science Advisory Secretariat]. 2010. Assessment of Georges Bank scallops (*Placopecten magellanicus*). Science Advisory Report 2010/036. 9 p.

Dickie LM, McCracken FD. 1955. Isopleth diagrams to predict equilibrium yield of a small flounder fishery. J. Fish. Res. Bd. Canada 12: 187-209.

Fairchild EA, Sulikowski JA, Rennels N, Howell WH, Tsang PCW. 2010. Effects of moving acclimation cages before release of cultured fish: alternative release strategies for a juvenile winter flounder *Pseudopleuronectes americanus* stock enhancement effort. Aquacult. Res. 41: 602-606.

Gavaris S. 1988. An Adaptive framework for the estimation of population size. CAFSAC Res. Doc. 88/29. 12pp.

Gavaris, S., G. Robert, and L. Van Eeckhaute. 2007. Discards of Atlantic Cod, Haddock, and Yellowtail Flounder from the 2005 and 2006 Canadian Scallop Fishery on Georges Bank. TRAC Reference Document 2007/03, 10p. http://www2.mar.dfo-mpo.gc.ca/science/TRAC/documents/TRD_2007_03_E.pdf

Gavaris S, Clark KJ, Hanke AR, Purchase CF, Gale J. 2010. Overview of discards from Canadian commercial fisheries in NAFO Divisions 4V, 4W, 4X, 5Y and 5Z for 2002-2006. Can. Tech. Rpt. Fish. Aquat. Sci. 2873. 104 p.

Gislason H, Daan N, Rice JC, Pope JG. 2010. Size, growth, temperature and the natural mortality of marine fish. Fish and Fisheries 11: 149-158

Greene C, Pershing A. 2003. The flip-side of the North Atlantic Oscillation and modal shifts in slope-water circulation patterns. Limnology and oceanography 48:319-322.

Hare JA, Able KW. 2007. Mechanistic links between climate and fisheries along the east coast of the United States: explaining population outbursts of Atlantic croaker (*Micropogonias undulatus*). Fisheries Oceanography 16:31-45.

Hare JA, Alexander M, Fogarty MJ, Williams E, Scott J. 2010. Forecasting the dynamics of a coastal fishery species using a coupled climate-population model. Ecological Applications 20:452-464.

Hare JA, Kane J. In press. Zooplankton of the Gulf of Maine – a changing perspective. American Fisheries Society Science Symposium.

Hettler WF, Chester AJ. 1982. The relationship of winter temperature and spring landings of pink shrimp, *Penaeus duorarum*, in North Carolina. Fishery Bulletin 80:761-768.

Hewitt DA, Hoenig JM. 2005. Comparison of two approaches for estimating natural mortality based on longevity. Fish. Bull. 103: 433-437.

Hoenig JM. 1983. Empirical use of longevity data to estimate mortality rates. Fish. Bull. 82: 898-903.

Hoenig JM, Barrowman NJ, Hearn WS, Pollock KH. 1998. Multiyear Tagging Studies Incorporating Fishing Effort Data. Can. J. Fish. Aquat. Sci. 55:1466-1476.

Howe AB, Coates PG. 1975. Winter flounder movements, growth and mortality off Massachusetts. Trans. Am. Fish. Soc. 104: 13-29.

Hurrell JW, Deser C. 2010. North Atlantic climate variability: the role of the North Atlantic Oscillation. Journal of Marine Systems 79:231-244.

Iles TC, Beverton RJH. 1998. Stock, recruitment and moderating processes in flatfish. Journal of Sea Research 39:41-55.

Joyce TM. 2002. One hundred plus years of wintertime climate variability in the eastern United States. Journal of Climate 15:1076-1086.

Joyce TM, Zhang R. 2010. On the Path of the Gulf Stream and the Atlantic Meridional Overturning Circulation. Journal of Climate 23:3146-3154.

Kalnay EC, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J. 1996. The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society **77**:437-471.

Keller AA, Klein-MacPhee G. 2000. Impact of elevated temperature on the growth, survival, and trophic dynamics of winter flounder larvae: a mesocosm study. Can. J. Fish. Aquat. Sci. 57:2382-2392.

Kerr RA. 2005. Atlantic climate pacemaker for millennia past, decades hence? Science 309:43-44.

Kendall WC. 1912. Notes on a new species of flatfish from off the coast of New England. Bulletin of the U.S. Bureau of Fisheries 30: 391-394.

Lazzari M. 1997. Monthly and annual means of sea surface temperature: Boothbay Harbor, Maine 1905–1996. Research Reference Document.

Levi D, Andreoli M, Bonanno A, Fiorentino F, Garofalo G, Mazzola S, Norrito G, Patti B, Pernice G, Ragonese S. 2003. Embedding sea surface temperature anomalies into the stock recruitment relationship of red mullet (Mullus barbatus L. 1758) in the Strait of Sicily. Scientia Marina 67:259-268.

Lux FE. 1973. Age and growth of the winter flounder, *Pseudopleuronectes americanus*, on Georges Bank. Fish. Bull. 71: 505-512.

Lux FE, Peterson AE Jr, Hutton RF. 1970. Geographical variation in fin ray number in winterflounder, *Pseudopleuronectes americanus* (Walbaum), off Massachusetts. Trans. Am. Fish. Soc. 1970(3): 483-488.

Manderson JP. 2008. The spatial scale of phase synchrony in winter flounder (*Pseudopleuronectes americanus*) production increased among southern New England nurseries in the 1990s. Can. J. Fish. Aquat. Sci. 65:340-351.

McBride RS, Wuenschel MJ, McElroy D, Rowinski Y, Thornton G, Nitschke P. MS 2011. Classifying winter flounder maturity during NEFSC resource surveys: comparing at-sea, macroscopic classifications with results from a gonad histology method. Stock Assessment Workshop Southern Demersal Working Group 52 Working Paper 8.

Miller TJ. Submitted. A hierarchical model for relative catch efficiency from gear selectivity and calibration studies.

Miller TJ. MS 2011. Winter flounder length-based survey calibration. Stock Assessment Workshop Southern Demersal Working Group 52 Working Paper 7.

Miller TJ, Das C, Politis PJ, Miller AS, Lucey SM, Legault CM, Brown RW, Rago PJ. 2010. Estimation of Albatross IV to Henry B. Bigelow calibration factors. NEFSC Ref Doc. 10-05. 233 p.

Myers RA, Bowen KG, Barrowman NJ. 1999. Maximum reproductive rate of fish at low population sizes. Can. J. Fish. Aquat. Sci. 56: 2404-2419.

NEFMC [New England Fishery Management Council]. 2009. Final Amendment 16 to the Northeast Multispecies Fishery Management Plan Including a Final Supplemental Environmental Impact Statement and an Initial Regulatory Flexibility Analysis. 905 pp. plus Appendices. http://www.nefmc.org/nemulti/planamen/Amend16/final%20amendment%2016/4.0_091016_Final_Amendment_16-3.pdf.

NEFSC. 1999. Report of the 28th Northeast Regional Stock Assessment Workshop (28th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. NEFSC Ref Doc. 99-08. 304 p.

NEFSC [Northeast Fisheries Science Center]. 2002a. Final report of the working group on re-evaluation of biological reference points for New England groundfish. 231 p. http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0204/

NEFSC [Northeast Fisheries Science Center]. 2002b. Report of the 34th Northeast Regional Stock Assessment Workshop (34th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. Northeast Fish. Sci. Cent. Ref. Doc. 02-06; 346 p.

NEFSC [Northeast Fisheries Science Center.] 2002c. Assessment of 20 Northeast groundfish stocks through 2001: a report of the Groundfish Assessment Review Meeting (GARM), Northeast Fisheries Science Center, Woods Hole, Massachusetts, October 8-11, 2002. Northeast Fish. Sci. Cent. Ref. Doc. 02-16; 521 p.

NEFSC. 2003. Report of the 36th Northeast Regional Stock Assessment Workshop (36th SAW): Stock Assessment Review Committee (SARC) consensus summary of assessments. NEFSC Ref Doc. 03-06. 453 p.

NEFSC. 2005. Assessment of 19 Northeast groundfish stocks through 2004. 2005 Groundfish Assessment Review Meeting (2005 GARM), Northeast Fisheries Science Center, Woods Hole, Massachusetts. 15-19 August 2005. NEFSC Ref Doc. 05-13. 508 p.

NEFSC. 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. NEFSC Ref Doc. 08-15. 884 p.

Northeast Fisheries Science Center. 2010. 50th Northeast Regional Stock Assessment Workshop (50th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 10-09; 57 p.

NEFSC. 2011. 51st Northeast regional stock assessment workshop (51st SAW) assessment report. NEFSC Ref. Doc. 11-02. 856 p.

NOAA Fisheries Toolbox (NFT). 2010. Virtual Population Analysis Model (VPA/ADAPT) version 3.0.3. Internet website http://nft.nefsc.noaa.gov/VPA.

Neill WH, Miller JM, Van Der Veer HW, Winemiller KO. 1994. Ecophysiology of marine fish recruitment: a conceptual framework for understanding interannual variability. Netherlands Journal of Sea Research 32:135-152.

Nixon SW, Granger S, Buckley BA, Lamont M, Rowell B. 2004. A one hundred and seventeen year coastal water temperature record from Woods Hole, Massachusetts. Estuaries 27:397-404.

Nye JA, Joyce TM, Kwon YO, Link JS. *In review*. Silver hake tracks changes in Northwest Atlantic circulation. Nature Communications.

Nye JA, Link JS, Hare JS, Overholtz WJ. 2009. Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. Mar. Ecol. Prog. 393: 111-129.

O'Boyle, R., M. Bell, S. Garvaris, V. Haist, S. Reeves, and G. Thompson. 2008. Panelist summary report of the groundfish assessment review meeting (GARM III), Part3. Biological reference points. 92p.
http://nefsc.noaa.gov/saw/garm/Garm%20III_BRPs_report_6june2008_finalCorrected.pdf

O'Brien L, Burnett J, Mayo R. 1993. Maturation of nineteen species of finfish off the northeast coast of the United States, 1985-1990. NOAA Technical Report NMFS 113. 66 pp.

Ottersen GB, Planque B, Belgrano A, Post E, Reid P, Stenseth N. 2001. Ecological effects of the North Atlantic Oscillation. Oecologia 128:1-14.

Roelofs EW, Bumpus DF. 1953. The hydrography of Pamlico Sound. Bulletin of Marine Science of the Gulf and Caribbean 3:181-205.

Palmer MC, Wigley SE. MS 2011. Using positional data from vessel monitoring systems (VMS) to validate the logbook-reported area fished and the stock allocation of commercial fisheries landings, 2004-2008. Stock Assessment Workshop Southern Demersal Working Group 52 Working Paper 9.

Perlmutter A. 1947. The blackback flounder and its fishery in New England and New York. Bulletin of the Bingham Oceanographic Collection, 11: 1- 92.

Pierce DE, Howe AB. 1977. A further study on winter flounder group identification off Massachusetts. Trans. Am. Fish. Soc. 106: 131-139.

Poole JC. 1969. A study of winter flounder mortality rates in Great South Bay, New York. Trans. Am. Fish. Soc. 98: 611-616.

Prager, M.H. 2004. User's manual for ASPIC: a stock production model incorporating covariates (ver.5). Beaufort Lab. Doc. BL-2004-01. 27 p.

Prager, M.H. 1995. User's manual for ASPIC: a stock production model incorporating covariates, program version 3.6x. Miami Lab. Doc. MIA-92/93-55. 25 p.

Quinn TJ II, Deriso RB. 1999. Quantitative fish dynamics. Oxford Univ. Press, New York, NY.

Rademeyer RA, Butterworth DS. MS 2010. Initial applications of statistical catch-at-age assessment methodology to the Southern New England/Mid-Atlantic winter flounder resource. 29 p.

Roelofs EW, Bumpus DF. 1953. The hydrography of Pamlico Sound. Bulletin of Marine Science of the Gulf and Caribbean 3:181-205.

Rothschild BJ. 1986. Dynamics of marine fish populations. Harvard University Press, Cambridge, Massachusetts.

Saila SB, Horton DB, Berry RJ. 1965. Estimates of the theoretical biomass of juvenile winter flounder, *Pseudopleuronectes americanus* (Walbaum) required for a fishery in Rhode Island. J. Fish. Res. Bd. Canada 22: 945-954.

Sameoto J. 2011. CA Division of Fisheries and Oceans. St. Andrews, New Brunswick, CA. pers. comm.

Sissenwine, M. P. and E. W. Bowman. 1978. An analysis of some factors affecting the catchability of fish by bottom trawls. ICNAF Res. Bull. No. 13. 7 p.

Smith WG. 1985. Temporal and spatial spawning patterns of the principal species of fish and invertebrates in the Georges Bank region. NMFS/NEFSC, Sandy Hook Laboratory, Lab. Ref. Doc. 85-4. 35 p.

Smith BE, Link JS. 2010. The trophic dynamics of 50 finfish and two squid species on the Northeast U.S. continental shelf. NOAA Tech. Memo. NMFS-NE-216. 640 p.

Stone, Heath. 2011. CA Division of Fisheries and Oceans. St. Andrews, New Brunswick, CA. pers. comm.

Stoner A, Manderson J, Pessutti J. 2001. Spatially explicit analysis of estuarine habitat for juvenile winter flounder combining generalized additive models and geographic information systems. Marine Ecology Progress Series 213:253–271.

Taylor AH, Stephens JA. 1998. The North Atlantic oscillation and the latitude of the Gulf Stream. Tellus A 50:134-142.

Taylor DL, Collie JS. 2003. Effect of temperature on the functional response and foraging behavior of the sand shrimp Crangon septemspinosa preying on juvenile winter flounder *Pseudopleuronectes americanus.* Marine Ecology Progress Series 263:217-234.

Terceiro M. MS 2011. Response to SNE/MA winter flounder TOR 4. Stock Assessment Workshop Southern Demersal Working Group 52 Working Paper 5.

Thompson WF, Bell FH. 1934. Biological statistics of the Pacific halibut fishery. 2. Effect of changes in intensity upon total yield and yield per recruit of gear. Rep Int Fish (Pacific halibut) Comm. 8: 49 p.

Vetter EF. 1988. Estimation of natural mortality in fish stocks: a review. Fish. Bull. 86: 25-43.

Van Eeckhaute, L. and J. Brodziak. 2005. Assessment of eastern Georges Bank haddock. Transboundary Resource Assessment Committee [TRAC] Ref. Doc. 2005/03: 73 p.

Visbeck M, Chassignet EP, Curry RG, Delworth TL, Dickson RR, Krahmann G. 2003. The ocean's response to North Atlantic Oscillation variability. Page 279 *in* JW Hurrel, Y Kushnir, G Ottersen, and M Visbeck, editors. The North Atlantic Oscillation: Climatic significance and environmental impact. American Geophysical Union, Washington, D.C.

Waring GT, Josephson E, Fairfield CP, Maze-Foley K. 2009. US Atlantic and Gulf of Mexico marine mammal stock assessments 2008. NOAA Tech. Memo. NMFS-NE-210. 429 p.

Witherell DB, Burnett J. 1993. Growth and maturation of winter flounder, *Pseudopleuronectes americanus,* in Massachusetts. Fish. Bull. 91: 816-820.

Wigley, SE, P. Hersey and J. E. Palmer 2008a. A description of the allocation procedure applied to the 1994-2007 commercial landings data. NEFSC Ref. Doc 08-18. 61 p.

Wigley SE, Palmer MC, Blaylock J, Rago PJ. 2008b. A brief description of the discard estimation for the National Bycatch Report. NEFSC Ref. Doc. 08-02. 35 p.

Wood A. MS 2011. Re-analysis of Howe and Coates (1975): winter flounder natural mortality derived from data in Howe and Coates (1975) using instantaneous rates tagging models. Stock Assessment Workshop Southern Demersal Working Group 52 Working Paper 15.

Wuenschel M, Able K, Byrne D. 2009. Seasonal patterns of winter flounder *Pseudopleuronectes americanus* abundance and reproductive condition on the New York Bight continental shelf. Journal of Fish Biology 74:1508-1524.

Acknowledgements

This assessment could not have been conducted without: the data preparation and technical help provided by Mark Terceiro, Susan Wigley, and Chris Legault; the cross plots provided by Michele Thompson; the age data provided by Jay Burnett and Grace Thompson; and the multitude of data collected by the NMFS port agents and and NEFSC staffs from DMS, NEFOP and ESB. Many thanks to Christa Waters, Lou Van Eeckhaute, and Heath Stone (CA Department of Fisheries and Oceans) for providing the Canadian survey, landings and discard data.