Draft Working Paper for Pre-Dissemination Peer Review Only

Working Paper-1 Assess Rpt
June 6, 2008

Summer Flounder Assessment for 2008

SARC 47
June 16-20, 2008

Assessment Report

by

SAW Southern Demersal Working Group
National Marine Fisheries Service
Northeast Fisheries Science Center
166 Water Street

SUMMARY

This June 2008 assessment of the summer flounder (*Paralichthys dentatus*) stock along the Atlantic coast (Maine to North Carolina) is an update through 2007 of commercial and recreational fishery catch data, research survey indices of abundance, and the analyses of the data. For 2007, commercial and recreational fishery final quotas were 4,401 mt and 3,030 mt, respectively, for a total of 7,431 mt. The reported commercial landings used in this assessment for 2007 were 4,489 mt, while estimated recreational landings were 4,445 mt, for a 2007 landings total of 8,934 mt. The 2008 commercial and recreational final quotas are 4,292 mt and 2,861 mt, respectively, for a total of 7,047 mt.

Three modeling approaches were explored in detail for this assessment. A virtual population analysis (VPA) of commercial and recreational total catch at age (landings plus discards) was conducted. In addition two statistical catch-at-age models were explored (ASAP and SS2). The same suites of fishery-independent indices of recruitment and stock abundance were used in all three modeling approaches; these were developed from the Northeast Fisheries Science Center (NEFSC) winter, spring, and autumn surveys; Massachusetts spring and autumn surveys; Rhode Island survey; Connecticut spring and autumn surveys; and New Jersey trawl survey. Recruitment indices from surveys conducted by the states of North Carolina, Virginia, and Maryland were also used. The WG identified the Age-structured Assessment Program (ASAP) as the best analytical tool to assess the summer flounder population.

The summer flounder stock is not overfished and overfishing is not occurring relative to the proposed 2008 assessment biological reference points. Fishing mortality calculated from the average of the currently fully recruited ages (3-7+) ranged between 1.143 and 2.042 during 1982-1996. The fishing mortality rate has declined to below 1.000 since 1996 and was estimated to be 0.288 in 2007, below the proposed fishing mortality reference point = F35% = FMSY = 0.310. There is an 80% probability that the fishing mortality rate in 2007 was between 0.253 and 0.325. Spawning stock biomass (SSB) declined from 24,674 mt in 1982 to 7,017 in 1989, then increased to 43,932 mt by 2004. SSB was estimated to be 43,363 in 2007, about 72% of the proposed SSB35% = SSBMSY reference point = 60,074 mt. There is an 80% chance that SSB in 2007 was between 39,325 and 48,122 mt. The arithmetic average recruitment from 1982 to 2007 is 41.6 million fish at age 0. The 1982 and 1983 year classes are the largest in the assessment time series, at 73.5 and 81.6 million fish; the 1988 year class is the smallest at only 12.8 million fish. The 2007 year class is currently estimated to be about 40.0 million fish.

The assessment exhibits a consistent retrospective pattern of underestimation of F and overestimation of SSB. There is no consistent retrospective pattern in recruitment evident. Over the last 3 years, the annual retrospective change in fishing mortality has ranged from +30 to -5%; over the last 3 years, the annual retrospective change in SSB has ranged from -29 to +6%. Future TALs that correspond to fishing at or near the fishing mortality rate threshold (F35% = FMSY = 0.310) may result in overfishing if the observed retrospective pattern persists. Therefore, managers should consider adopting

future TALs lower than those indicated by forecast median values to decrease the chance that overfishing will occur. If landings in 2008 are 7,153 mt (15.8 million lbs) and discards are 885 mt (2.0 million lbs), the forecast estimates a median (50% probability) F in 2008 = 0.238 and a median SSB on November 1, 2008 of 46,992 mt, above the proposed biomass threshold of one-half SSBMSY = 30,037 mt Fishing at Frebuild = 0.274 in 2009 results in forecast median (50%ile) landings of 9,211 mt (20.3 million lbs); the corresponding 25%ile of landings is 8,653 mt (19.1 million lbs). Continued fishing at Frebuild = 0.274 during 2010-2012 is forecast to rebuild the stock to SSBMSY = 60,074 in 2012.

TABLE OF CONTENTS

INTRODUCTION AND BACKGROUND

WORKING GROUP PROCESS

The Southern Demersal Working Group (WG) began work in 2007 to produce this benchmark assessment of summer flounder through 2007/2008. The WG met via conference call on November 2007 and in person in February 2008, April 2008, and May 2008. The Stock Assessment Workshop (SAW) was held during June 16-20, 2008 to present the benchmark assessment of summer flounder through 2007/2008 to the Stock Assessment Review Committee (SARC). The following scientists and managers participated in the WG meetings and benchmark assessment update:

Ken Able	Rutgers University
Jeff Brust	New Jersey Division of Fish and Wildlife (NJDFW)
Paul Caruso	Massachusetts Division of Marine Fisheries (MADMF)
Jessica Coakley	Mid-Atlantic Fishery Management Council (MAFMC)
Victor Crecco	Connecticut Department of Environmental Protection (CTDEP)
Greg DiDomenico	Garden State Seafood Association (GSSA)
Bruce Freeman	Partnership for Mid-Atlantic Fisheries Science (PMAFS)
Emerson Hasbrouck	Cornell University
Toni Kerns	Atlantic States Marine Fisheries Commission (ASMFC)
Laura Lee	Virginia Marine Resources Commission (VMRC)
Chris Legault	NMFS NEFSC
Brian Murphy	Rhode Island Department of Environmental Management, Division of Fish and Wildlife (RIDFW)
Mark Maunder	Quantitative Resource Assessment (QRA)
Paul Nitschke	NMFS NEFSC
Bill Overholtz	NMFS NEFSC
Eric Powell	Rutgers University
Paul Rago	NMFS NEFSC
Michael Ruccio	NMFS Northeast Regional Office (NMFS NERO)
Kathy Sosebee	NMFS NEFSC
Mark Terceiro	NMFS NEFSC
Alice Weber	New York Department of Environmental Conservation (NYDEC)
Greg Wojcik	Connecticut Department of Environmental Protection (CTDEP)
Richard Wong	Delaware Department of Fish and Wildlife (DEDFW)

Although they were unable to attend the meeting, David Simpson of the CTDEP, Don Byrne of the NJDFW, Stew Michels of the DEDFW, Steve Doctor of the Maryland Department of Natural Resources (MDDNR), Chris Bonzak of the Virginia Institute of Marine Science (VIMS), Rob O'Reilly of the VMRC, and Chris Batsavage of the North Carolina Division of Marine Fisheries (NCDMF) provided research survey and/or fisheries catch data used in the assessment.

STOCK UNIT

For assessment purposes, the previous definition of Wilk et al. (1980) of a unit stock extending from Cape Hatteras north to New England has been accepted in this and previous assessments (e.g., NEFSC 2002). The Mid-Atlantic Fishery Management Council (MAFMC) and the Atlantic States Marine Fisheries Commission (ASMFC) Fishery Management Plan (FMP) define the management unit for summer flounder as extending from the southern border of North Carolina, northward to the U.S.-Canadian border. A recent summer flounder genetics study, which revealed no population subdivision at Cape Hatteras (Jones and Quattro, 1999), is consistent with the definition of the management unit. Recent consideration of summer flounder stock structure incorporating new tagging data concluded that evidence supported the existence of stocks north and south of Cape Hatteras, with the stock north of Cape Hatteras possibly composed of two distinct spawning aggregations, off New Jersey and Virginia-North Carolina (Kraus and Musick, 2003). The conclusions of Kraus and Musick (2003) are consistent with the current assessment stock unit.

HISTORY OF MANAGEMENT AND THE ASSESSMENT

An overview of the history of the summer flounder FMP and assessment is provided in this section and the box below. Management of the summer flounder fishery began through the implementation of the original Summer Flounder FMP, which was approved by National Marine Fisheries Service in 1988. This 1988-1989 time period coincides with the lowest levels of stock biomass for summer flounder since 1982.

There are two management entities that cooperatively develop fishery regulations for this resource; the ASMFC and the MAFMC in conjunction with the National Marine Fisheries Service (NMFS) as the federal implementation and enforcement entity. The cooperative management endeavor was developed because a significant portion of the catch is taken from both state (0-3 miles offshore) and federal waters (3-200 miles offshore).

Amendment 1 to the FMP in 1990 established the fishing definition for summer flounder; it is the fishing mortality rate equal to F_{MAX}, initially estimated as 0.23 (NEFSC 1990). Amendment 2 (1992) established target fishing mortality rates for summer flounder for 1993-1995 as F = 0.53, and F_{MAX} = 0.23 for 1996 and beyond. Regulations enacted under Amendment 2 to meet those fishing mortality rate targets included: 1) an annual fishery landings quota, with 60% allocated to the commercial fishery and 40% to the recreational fishery, based on the historical (1980-1989) division of landings; the commercial allocation is further distributed among the states based on their share of commercial landings during 1980-1989; 2) commercial minimum landed fish size limit at 13 in (33 cm), as established in the original FMP; 3) a minimum mesh size of 5.5 in (140 mm) diamond or 6.0 in (152 mm) square for commercial vessels using otter trawls that possess 100 lbs (45 kg) or more of summer flounder, with exemptions for the flynet fishery and vessels fishing in an exempted area off southern New England (the Northeast Exemption Area) during 1 November to 30 April; 4) permit requirements for the sale and purchase of summer flounder; and 5) annually adjustable regulations for the recreational fishery, including seasons, a 14 in (36 cm) minimum landed fish size, and possession limits.

The results of previous assessments indicated that summer flounder abundance was not increasing as rapidly as projected when Amendment 2 regulations were implemented. In anticipation of the need to drastically reduce fishery quotas in 1996 to meet the management target of F_{MAX} , the MAFMC and ASMFC modified the fishing mortality rate reduction schedule in 1995 to allow for more stable landings from between years, while slowing the rate of stock rebuilding. Amendment 7 to the FMP set target fishing mortality rates of 0.41 for 1996 and 0.30 for 1997, with a target of F_{MAX} = 0.23 for 1998 and beyond. Total landings were to be capped at 8,400 mt (18.51 million lbs) in 1996-1997, unless a higher quota in those years provided a realized F = 0.23.

Amendment 12 (1999) defined overfishing for summer flounder as occurring when the fishing mortality rate exceeds the threshold fishing mortality rate of F_{MSY}. Because F_{MSY} could not be reliably estimated for summer flounder, F_{MAX} = 0.24 was used as a proxy for F_{MSY}. This was also defined as the target fishing mortality rate (F_{TARGET}= F_{MSY}= F_{MAX}). Under this amendment, the stock was defined to be overfished when total stock biomass falls below the minimum biomass threshold of one-half of the biomass target, B_{MSY}. Because B_{MSY} could not be reliably estimated, the biomass target was defined as the product of total biomass per recruit and contemporary (1982-1996) median recruitment, at that time estimated to be 153,350 mt (338 million lbs), with the minimum biomass threshold defined as 76,650 mt (169 million lbs). Through the 1999 stock assessment (Terceiro 1999), those reference points were updated using updates of median recruitment (1982-1998) and mean weights at age (1997-1998), which resulted in a biomass target of 106,444 mt (235 million lbs) and minimum biomass threshold of 53,222 mt (118 million lbs). The Terceiro (1999) reference points were retained in the 2000 and 2001 stock assessments (NEFSC 2000, MAFMC 2001a) because of the stability of the input data. Concurrent with the development of the 2001 assessment, the MAFMC and ASMFC convened the ASMFC Summer Flounder Overfishing Definition Review Committee to review these biological reference points. The work of this Committee was later reviewed by the MAFMC Scientific and Statistical Committee (SSC) in August 2001. The SSC recommended that using the F_{MSY} proxy for F_{MAX} = 0.26 was appropriate and be retained for 2002, and endorsed the recommendation of SARC 31 (NEFSC 2000) which stated that "...the use of F_{MAX} as a proxy for F_{MSY} should be reconsidered as more information on the dynamics of growth in relation to biomass and the shape of the stock recruitment function become available" (MAFMC 2001b).

The benchmark stock assessment in 2002 (SAW 35; NEFSC 2002) indicated the summer flounder stock was overfished and overfishing was occurring relative to the current biological reference points. The fishing mortality rate had declined from 1.32 in 1994 to 0.27 in 2001, marginally above the overfishing reference point ($F_{THRESHOLD}$ = F_{TARGET} = F_{MAX} = 0.26). Total stock biomass in 2001 was estimated as 42,900 mt (94.6 million lbs), or 19% below the biomass threshold (53,200 mt; 117.3 million lbs). The review of the 2002 stock assessment (SARC 35) concluded that updating the biological reference points was not warranted at that time (NEFSC 2002). Updates to the stock assessment were completed in 2003 (Terceiro 2003), 2004 (SDWG 2004), and 2005 (SAW 41;

NEFSC 2005). While the 2003 assessment found the summer flounder stock was not overfished and no overfishing was occurring, the 2004 and 2005 assessments found the stock again experiencing overfishing. The 2005 SAW 41 assessment recommended updating the values for the fishing mortality and stock biomass reference points.

The most recent assessment peer review on summer flounder was the NMFS Office of Science and Technology Division (S&T) Peer Review of the 2006 SDWG assessment (October 2006; Terceiro 2006a, 2006b). This review made several recommendations, including modification of the definition of the overfished stock from what was originally defined under Amendment 2 to the FMP. Instead of using total stock biomass (as estimated on January 1), the stock was now considered overfished when November 1 spawning stock biomass fell below one-half SSBMSY = 44,706 mt (98.6 million lbs). The 2007 assessment update (SDWG 2007) found that relative to the biological reference points, the stock is overfished and overfishing is occurring. The fishing mortality rate estimated for 2006 was 0.35, a significant decline from the 1.32 estimated for 1994 but above the threshold F of 0.28. The assessment presented in this document builds off the recommendations of numerous peer reviews since the FMP was implemented that have facilitated methodological improvements in the assessment and reference point calculations.

Summary of the history of the Summer Flounder, Scup, and Black Sea Bass FMP.

Year	Document	Plan Species	Management Action
1988	Original FMP	summer flounder	- Established management plan for summer flounder
1991	Amendment 1	summer flounder	- Established an overfishing definition for summer flounder
1993	Amendment 2	summer flounder	- Established rebuilding schedule, commercial quotas, recreational harvest limits, size limits, gear restrictions, permits, and reporting requirements for summer flounder - Created the Summer Flounder Monitoring Committee
1993	Amendment 3	summer flounder	- Revised the exempted fishery line - Increased the large mesh net threshold - Established otter trawl retentions requirements for large mesh use
1993	Amendment 4	summer flounder	- Revised state-specific shares for summer flounder quota allocation
1993	Amendment 5	summer flounder	- Allowed states to combine or transfer commercial summer flounder quota
1994	Amendment 6	summer flounder	- Set criteria for allowance of multiple nets on board commercial vessels for summer flounder - Established deadline for publishing catch limits, commercial mgmt. measures for summer flounder
1995	Amendment 7	summer flounder	- Revised the F reduction schedule for summer flounder

Summary of the history of the Summer Flounder, Scup, and Black Sea Bass FMP.			
Year	**Document**	**Plan Species**	**Management Action**
1996	Amendment 8	summer flounder and scup	- Incorporated Scup FMP into Summer Flounder FMP and established scup measures including commercial quotas, recreational harvest limits, size limits, gear restrictions, permits, and reporting requirements
1996	Amendment 9	summer flounder and black sea bass	- Incorporated Black Sea Bass FMP into Summer Flounder FMP and established black sea bass measures including commercial quotas, recreational harvest limits, size limits, gear restrictions, permits, and reporting requirements
1997	Amendment 10	summer flounder, scup, and black sea bass	- Modified commercial minimum mesh requirements, continued commercial vessel moratorium, prohibited transfer of fish at sea, and established special permit for party/charter sector for summer flounder
1998	Amendment 11	summer flounder, scup, and black sea bass	- Modified certain provisions related to vessel replacement and upgrading, permit history transfer, splitting, and permit renewal regulations
1999	Amendment 12	summer flounder, scup, and black sea bass	- Revised FMP to comply with the SFA and established framework adjustment process
2001	Framework 1	summer flounder, scup, and black sea bass	-Established quota set-aside for research for all three species
2001	Framework 2	summer flounder	- Established state-specific conservation equivalency measures for summer flounder
2003	Amendment 13	summer flounder, scup, and black sea bass	- Addressed disapproved sections of Amendment 12 and included new EIS
2003	Framework 3	scup	- Allowed the rollover of winter scup quota - Revised start date for summer quota period for scup fishery
2003	Framework 4	scup	- Established system to transfer scup at sea
2004	Framework 5	summer flounder, scup, and black sea bass	- Established multi-year specification setting of quota for all three species
2006	Framework 6	summer flounder	- Established region-specific conservation equivalency measures for summer flounder
2007	Amendment 14	scup	- Established rebuilding schedule for scup
2007	Framework 7	summer flounder, scup, and black sea bass	- Built flexibility into process to define and update status determination criteria for each plan species - Scup GRAs made modifiable through framework adjustment process

TERMS OF REFERENCE

1.0 Characterize the commercial and recreational catch, effort and CPUE, including descriptions of landings, discards and discard mortality.

1.1 Commercial Fishery Landings

Total U.S. commercial landings of summer flounder from Maine to North Carolina peaked in 1979 at nearly 18,000 mt (39.7 million lbs, Table 1, Figure 1). The reported landings in 2007 of 4,489 mt (9.89 million lbs) were slightly over the final 2007 quota of 4,401 mt (9.79 million lbs). Since 1980, about 70% of the commercial landings of summer flounder have come from the Exclusive Economic Zone (EEZ; greater than 3 miles from shore). Large variability in summer flounder landings exist among the states, over time, and the percent of total summer flounder landings taken from the EEZ has varied widely among the states.

Northeast Region (Maine to Virginia)

Annual commercial landings data for summer flounder in years prior to 1994 were obtained from trip-level detailed landings records contained in master data files maintained by the NEFSC (the "weighout system"; 1963-1993) and from summary reports of the Bureau of Commercial Fisheries and its predecessor the U.S. Fish Commission (1940-1962). Beginning in 1994, landings estimates were derived from mandatory dealer reports under the current NMFS Northeast Region (NER) summer flounder quota monitoring system.

Prior to 1994, summer flounder commercial landings were allocated to NEFSC 3-digit statistical area according to interview data (Burns et al. 1983). During 1994-2007, dealer landings were allocated to statistical area using fishing Dealer and fishing Vessel Trip Reports (VTR data) in a multi-tiered allocation procedure at the fishing-trip level (Wigley et al., 2007) A comparison of the distribution of landings by state and month as indicated by the dealer, VTR, and exact matched set data for trips with summer flounder landings during 1994-2007 is presented in Tables 2-15. Since the implementation of the annual commercial landings quota in 1993, the commercial landings have become concentrated during the first calender quarter of the year, with about 50% of the landings taken during the first quarter.

The distribution of Northeast Region (ME to VA) 1992-2007 landings by three-digit statistical area are presented in Table 16. Areas 537-539 (Southern New England), areas 611-616 (New York Bight), areas 621, 622, 625, and 626 (Delmarva region), and areas 631 and 632 (Norfolk Canyon area) have generally accounted for over 80% of the NER commercial landings. A summary of length and age sampling of summer flounder landings collected by the NEFSC commercial fishery port agent system in the NER is presented in Table 17. For comparability with the manner in which length frequency sampling in the recreational fishery has been evaluated, sampling intensity is expressed in terms of metric tons of landings (mt) per 100 fish lengths measured. The sampling is proportionally stratified by market category (jumbo, large, medium, small, and unclassified), with the sampling distribution generally reflecting the distribution of commercial landings by market category. Overall sampling intensity has improved markedly since 1995, from 165 mt per 100 lengths to 17 mt per 100 lengths (Table 17), and temporal and geographic coverage has generally improved as well (Tables 18-31).

The age composition of the NER commercial landings for 1994-2002 was generally estimated semiannually by market category and (usually) 1-digit statistical area (e.g., area 5 or area 6), using standard NEFSC procedures (market category length frequency samples converted to mean weights by length-weight relationships; mean weights in turn divided into landings to calculate numbers landed by market category; market category numbers at length apportioned to age by application of age-length keys, on semiannual area basis). For 2000-2002, sampling was generally sufficient to make quarterly estimates of the age composition in area 6 (in some cases, by division) for the large and medium market categories. For 2003-2007, sampling was generally sufficient to make quarterly estimates of the age composition in areas 5 and 6 for the jumbo, large, and medium market categories. The distribution of 1994-2007 length frequency samples by market category, 1- and 2-digit statistical area (division), and calendar quarter is presented in Tables 18-31.

NER landed numbers at age were raised to total NER (general canvas) commercial landings when necessary by assuming that landings not accounted for in the weighout/mandatory reporting system had the same age composition as that sampled. This was done as follows: calculate proportion at age by weight; apply proportions at age by weight to total NER commercial landings to derive total NER commercial catch at age by weight; divide by mean weights at age to derive total NER commercial landed numbers at age. The proportion of large and jumbo market category fish (generally of ages 3 and older) in the NER landings has increased since 1996, while the proportion of small market category landings (generally of ages 0 and 1) has become very low (Table 32, Figure 2). The mean size of fish landed in the NER commercial fishery has been increasing since 1993, and was 0.9-1.0 kg (2.0-2.2 lbs) during 2000-2007, typical of an age 3 summer flounder (Table 33).

North Carolina

The North Carolina winter trawl fishery accounts for about 99% of summer flounder commercial landings in North Carolina. A separate landings at age matrix for this component of the commercial fishery was developed from North Carolina Division of Marine Fisheries (NCDMF) length and age frequency sampling data. The NCDMF program samples about 10% of the winter trawl fishery landings annually, most recently (2005, 2006, and 2007) at a mean rate of 9 mt, 9 mt, and 5 mt of landings per 100 lengths measured, respectively (Table 34). All length frequency data used in construction of the North Carolina winter trawl fishery landings at age matrix were collected in the NCDMF program; age-length keys from NEFSC commercial data and NEFSC spring survey data (1982-1987) and NCDMF commercial fishery data (1988-2007) were combined by appropriate statistical area and semiannual period to resolve lengths to age. Fishery regulations in North Carolina also changed between 1987 and 1988, with increases in both the minimum mesh size of the codend and minimum landed fish size taking effect. It is not clear whether the change in regulations or the change in keys, or some combination, is responsible for the decreases in the numbers of age-0 and age-1 fish estimated in the North Carolina commercial fishery landings since 1987. Landed numbers at age (Figure 3) and mean weights at age from this fishery are shown in Tables 35-36.

1.2 Commercial Fishery Discards and Discard Mortality

In the 1993 SAW 16 assessment, analysis of variance of fishery observer data for summer flounder was used to identify stratification variables for an expansion procedure to estimate total landings and discards from the observer data kept and discard rates (weight per day fished) in the commercial fishery. Initial models included year, quarter, fisheries statistical division (2-digit area), area (divisions north and south of Delaware Bay), and tonnage class as main effects. Quarter and division consistently emerged as significant main effects without significant interaction with the year (NEFSC 1993). The estimation procedure expands transformation bias-corrected geometric mean catch (landings and discards) rates in year, quarter, and division strata by total days fished (days fished on trips landing any summer flounder by any mobile gear, including fish trawls and scallop dredges) to derive fishery landings and discards. The use of fishery effort as the multiplier (raising factor) allows estimation of landings from the fishery observer data for comparison with dealer reported landings, to help judge the potential accuracy of the procedure and/or sample data.

For strata with no fishery observer sampling, catch rates from adjacent or comparable strata were substituted as appropriate (except for Division 51, which generally has very low catch rates and negligible catch). Estimates of discard were stratified by 2 gear types (scallop dredges; trawls) for years when data were adequate (1992 and later years). Estimates at length and age were stratified by gear for 1994-2000 and 2002-2007, again due to sample size considerations. Only 11 fish were sampled from the sea scallop dredge fishery 2001, and so the scallop dredge discards were assumed to have the same length and age composition as the trawl fishery discards in 2001.

While estimates of catch rates from the NER fishery observer data were used in this assessment to estimate total discards, catch rate information is also reported in the VTR data. A comparison of discard to total catch ratios for the fishery observer and VTR data sets for trawl and scallop dredge gear indicates similar discard rates from the two data sources. Overall fishery observer and VTR discard to total catch ratios for 1994-2007 were generally within 10-15% of each other; 2001 was an exception, with an overall discard to total catch ratio of 49% in the fishery observer data and 29% in the VTR data. The most recent year (2007) was also an exception with an overall discard to total catch ratio of 59% in the fishery observer data and 36% in the VTR data. Discard rates of summer flounder in the scallop dredge fishery were much higher than in the trawl fishery (Tables 37-38).

The change in mid-1994 from the interview/weighout data reporting system to the VTR/mandatory dealer report system required a change in the estimation of effort (days fished) to estimate total discards. An initial examination of days fished and catch per unit effort (CPUE; landings per day fished) for cod conducted at SAW 24 (NEFSC 1997a) compared these quantities as reported in the full weighout and VTR data sets (DeLong et al., 1997). This comparison indicated a shift to a higher frequency of short trips (trips with one or two days fished reported), and to a mode at a lower rate of CPUE. It was not clear at SAW 24 if these changes were due to the change in reporting system (units reported not comparable), or real changes in the fishery, and so effort data reported by the VTR system were not used quantitatively in the SAW 24 assessments. In the SAW 25 assessment for summer flounder (NEFSC 1997b), a slightly different comparison was made. The port agent interview data for 1991-93 and merged dealer/VTR data for 1994-1996 (the matched set data), which under each system serve as the "sample" to characterize the total commercial landings, were compared in relative terms (percent frequency). For summer flounder, the percent frequency of short trips (lower number of days fished per trip) increased during 1991-1996, but not to the degree observed for cod, and the mode of CPUE rates for summer flounder increased in spite of lower effort per trip. For the summer flounder fishery, these may reflect actual changes in the fishery, due to increased restrictions on allowable landings per trip (trip landings limits might lead to shorter trips) and stock size increases (higher CPUE). As for cod, however, the influence of each of these changes (reporting system, management changes, stock size changes) has not been quantified. Total days fished in the summer flounder fishery were comparable between the period from 1989-1993 and 1994 (Tables 39-45; WO DF and WO/VTR DF). Since 1994, total days fished have ranged from 20,670 days in 1999 to 8,872 days in 2007, with a mean of about 12,000 days, a substantial decline relative to the 1989-1993 mean of 22,000 days (Tables 46-73). Questions will remain about the accuracy of the VTR data. However, because the effort measure is critical to the estimation of discards for summer flounder, the VTR data were used as the best data source to estimate summer flounder fishery days fished for 1994-2007.

Two adjustments were made to the dealer/VTR matched data subset days fished estimates to fully account for summer flounder fishery effort during 1994-2007. First, the landings to days fished relationship in the matched set was assumed to be the same for unmatched trips, and so the days fished total in each discard estimation stratum (2-digit area and quarter) was raised by the dealer to matched set landings ratio. This step in the estimation accounted for days fished associated with trips landing summer flounder, and provided an estimate of discard for trips landing summer flounder (Tables 46-73, variable OB EST DISC 1).

Given the restrictions on the fishery however, there is fishing activity which results in summer flounder discards, but no landings, especially in the scallop dredge fishery. The days fished associated with these trips was accounted for by raising strata discard estimates by the ratio of the total days fished on trips catching any summer flounder (trips with landings and discard, plus trips with discard only) to the days fished on trips landing summer flounder (trips with landings and discard) (Tables 46-73, variable NO KEPT RATIO), for VTR trips reporting discard of any species (DeLong et al. 1997). For this step , it is necessary to assume that the discard rate (as indicated by the fishery observer data, which includes trips with discard but no landings, and which is used in previous estimation procedure steps) is the same for trips with only discards as for trips with both landings and discards.

Discard estimates for 1989-2007 are summarized in Tables 39-73 (variable OB EST DISC MT). Commercial fishery discard mortality in weight was highest in 1990-1991 and 1999, and lowest in 2004-2005 (Table 74). Estimates of landings from observer data ranged from +53% (1999) to -77% (2007) of the reported landings in the fisheries (Table 75), with discards ranging from 38% (1990) to 6% (1995) of the dealer reported landings. Total discards estimated for 2005, 2006, and 2007 were 10%, 10%, and 16% of the reported landings. Scallop dredge fishery discard to landed ratios are much higher than trawl fishery ratios, purportedly because of closures and trip limits. Although the scallop dredge landings of summer flounder are less than 5% of the total, the discards of summer flounder are of the same order of magnitude as in the trawl fishery.

The discard estimates described above were based only on the day fished data for ports in the NER during 1989-1996, and so it was necessary to raise the discard estimate to account for discarding occurring outside the NER reporting system (i.e., NER state reporting systems such as Connecticut and Virginia, and North Carolina). To determine the proper raising factor, landings accounted for by the NER reporting system (which result from the fishing effort on which the fishery observer discard estimate is based) were compared with total NER landings, plus that portion of North Carolina landings from the EEZ (it is assumed that only the North Carolina fishery in the EEZ would experience significant discard, as mesh regulations in state waters have resulted in very low discards in state waters since implementation of the regulation in 1989; R. Monaghan, NCDMF; personal communication, June 30, 1997). As a result of this exercise, the total discard estimates were raised by 11 to 38% for the 1989-1996 period. Since 1996, all states' landings and are included in the NER dealer reporting system, so no raising is necessary to account for missing landings. As recommended by SAW 16 (NEFSC 1993), a commercial fishery discard mortality rate of 80% was assumed to develop the final estimate of discard mortality (Table 66). The group did consider some preliminary information from a 2007 Cornell University Cooperative Extension study which conducted ten scientific trips that were made on inshore multispecies commercial trawling vessels to determine discard mortality rates relative to tow duration, fish size, and the amount of time fish were on the deck of the vessel (Working Paper 2; Appendix 1). The median mortality for all tows combined was 78.7%, very close to the estimated overall discard mortality of 80% currently used in the summer founder assessment. The mean of 64.6% however is considerably less. The SDWG recommended additional work be conducted to understand factors affecting discard mortality rates and the difference between the inshore (day-trip) and offshore (multi-day) components of the multispecies trawl fishery to facilitate future application of this information at a broader scale.

Existing fishery observer data were used to develop estimates of commercial fishery discard for 1989-2007. However, adequate data (e.g., interviewed trip data, survey data) are not available to develop summer flounder discard estimates for 1982-1988. Discard numbers were assumed to be very small relative to landings during 1982-1988 (because of the lack of a minimum size limit in the EEZ), but to have increased since 1989 with the implementation of fishery regulations under the FMP. It was recognized that not accounting directly for commercial fishery discards in 1982-1988 would result in an underestimation of fishing mortality and population sizes in these years.

NEFSC fishery observer length frequency samples were converted to sample numbers at age and sample weight at age frequencies by application of NEFSC survey length-weight relationships and fishery observer, commercial fishery, and survey age-length keys. Sample weight proportions at age were next applied to the raised fishery discard estimates to derive fishery total discard weight at age. Fishery discard weights at age were then divided by fishery observer mean weights at age to derive fishery discard numbers at age. Classification to age for 1989-1993 was done by semiannual (quarters 1 and 2 pooled, quarters 3 and 4 pooled) periods using NEFSC fishery observer age-length keys, except for 1989, when first period lengths were aged using combined commercial landings (quarters 1 and 2) and NEFSC spring survey age-length keys. For 1994-2007, only NEFSC winter, spring, and fall survey age-length keys were used, since fishery observer age-length keys were not yet available and commercial landings age-length keys contained an insufficient number of small summer flounder (<40 cm = 16 inches) that comprise most of the discards. Fishery observer sampling intensity and estimates are summarized in Table 74. Table 75 presents a comparison of commercial fishery dealer reported landings of summer flounder with estimates of summer flounder commercial landings from landings rates of NEFSC Domestic Observer sampling and commercial fishing effort (days fished) reported on commercial Vessel Trip Reports (VTR). Estimates of discarded numbers at age, mean length and mean weight at age are summarized in Tables 76-78.

The reason for discarding in the trawl and scallop dredge fisheries has been changing over time. During 1989 to 1995, the minimum size regulation was recorded as the reason for discarding summer flounder in over 90% of the observed trawl and scallop dredge tows. In 1999, the minimum size regulation was provided as the reason for discarding in 61% of the observed trawl tows, with quota or trip limits given as the discard reason in 26% of the observed tows, and high-grading in 11% of the observed tows. In the scallop fishery in 1999, quota or trip limits was given as the discard reason in over 90% of the observed tows. During 2000-2005, minimum size regulations were identified as the discard reason in 40-45% of the observed trawl tows, quota or trip limits in 25-30% of the tows, and high grading in 3-8%. In the scallop fishery during 2000-2005, quota or trip limits was given as the discard reason for over 99% of the observed tows. During 2006-2007, minimum size regulations were identified as the discard reason in 15-20% of the observed trawl tows, quota or trip limits in 60-70% of the tows, and high grading in 5-10%. In the scallop fishery during 2006-2007, quota or trip limits was given as the discard reason for about 40% of the observed tows, with about 50% reported as "unknown." As a result of the increasing impact of trip limits, fishery closures, and high grading as reasons for discarding, the age structure of the summer flounder discards has also changed, with a higher proportion of older fish being discarded (Table 76, Figures 4 and 5).

The WG considered other methods for the calculation of the discard estimates (Working Paper 1; Appendix 1); but ultimately determined the current methods are appropriate for the current assessment (i.e. make no changes to the discard estimation approach used in the 2008 benchmark assessment). It was recommended, however, in the working paper that future work focus on trawl and scallop dredge gear; other approaches should be examined such as using sums of ratio (NBRD2) estimators with alternative landings or effort raising factors, possibly for a "characteristic" group of landed species trips in the trawl fishery (e.g., fluke, scup, black sea bass, *Loligo* and *Illex* squids, yellowtail flounder, winter flounder, cod, haddock, silver hake, etc.)

1.3 Recreational Fishery Landings

Summary landings statistics for the summer flounder recreational fishery (catch type A+B1) as estimated by the National Marine Fisheries Service (NMFS) Marine Recreational Fishery Statistics Survey (MRFSS) are presented in Tables 79 and 80. Recreational fishery landings decreased 19% by number and 12% by weight from 2006 to 2007, as the fishery landed 47% over (4,445 mt; 9.80 million lbs) the harvest limit established for 2007 of 3,030 mt (6.68 million lbs).

The commercial fishery VTR system provides an alternative set of reported recreational landings by the party/charter boat sector. A comparison of VTR reports and MRFSS estimates indicates that MRFSS estimates are higher by an average factor of 2.68 for the 1995-2007 period, with an increasing trend in recent years and ranging from a factor of 1.02 in 1998 to 5.47 in 2005 (Table 81). It is unclear if this is due mainly to under-reporting of party/charter boat recreational landings in the VTR system, or a systematic positive bias of MRFSS landings estimates for the party/charter boat sector.

Length frequency sampling intensity for the recreational fishery for summer flounder was calculated by MRFSS subregions (North - Maine to Connecticut; Mid - New York to Virginia; South - North Carolina) based on a metric tons of landings per hundred lengths measured basis (Burns et al. In Doubleday and Rivard, 1983). For 2007, aggregate sampling intensity averaged 132 mt of landings per 100 fish measured (Table 82).

MRFSS sample length frequency data, NEFSC commercial age-length data, and NEFSC survey age-length data were examined in terms of number of fish measured/aged on various temporal and geographical bases. Correspondences were made between MRFSS intercept date (quarter), commercial quarter, and survey season (spring and summer/fall), and between MRFSS subregion, commercial statistical areas, and survey depth strata to integrate data from the different sources. Based on the number, size range, and distribution of lengths and ages, a semiannual (quarters 1 and 2; quarters 3 and 4), subregional basis of aggregation was adopted for matching of commercial and survey age-length keys with recreational length frequency distributions to convert lengths to ages.

Limited MRFSS length sampling for larger fish resulted in a high degree of variability in mean length for older fish, especially at ages 5 and older during the first decade of the time series. Attempts to estimate length-weight relationships from the MRFSS biological sampling data provided unsatisfactory results. As a result, quarterly length (mm) to weight (g) relationships from Lux and Porter (1966) were used to calculate annual mean weights at age from the estimated age-length frequency distribution of the landings.

The recreational landings historically were dominated by relatively young fish. Over the 1982-1996 period, age 1 fish accounted for over 50% of the landings by number; summer flounder of ages 0 to 4 accounted for over 99% of landings by number. No fish from the recreational landings were determined to be older than age 7. With increases in the minimum size since 1996 (to 14.5 in [37 cm] in 1997, 15 in [38 cm] in 1998-1999, generally 15.5 in [39 cm] in 2000, and various state minimum sizes from 14.0 [36 cm] to 19.5 in [50 cm] in 2001-2007) and a trend to lower fishing mortality rates, the age composition of the recreational landings now includes mainly fish at ages 3 and 4. The number of summer flounder of ages 4 and older landed by the recreational fishery in 2007 (34% of the landings by number) was the highest in the time series (Table 83, Figure 6).

1.4 Recreational Fishery Discards and Discard Mortality

MRFSS catch estimates were aggregated on a subregional basis for calculation of the proportion of live discard (catch type B2) to total catch (catch types A+B1+B2) in the recreational fishery for summer flounder. The live discard has varied from about 18% (1985) to about 86% (2007) of the total catch during 1982-2007 (Table 84).

To account for all removals from the summer flounder stock by the recreational fishery, some assumptions about the biological characteristics and hooking mortality rate of the recreational live discard need to be made, because biological samples are not routinely taken of MRFSS catch type B2 fish. In previous assessments, data available from New York Department of Environmental Conservation (NYDEC) surveys (1988-92) of New York party boats suggested the following: 1) nearly all (>95%) of the fish released alive from boats were below the minimum regulated size (during 1988-92, 14 in [36 cm] in New York state waters); 2) nearly all of these fish were age 0 and age 1 summer flounder; and 3) age 0 and 1 summer flounder occurred in approximately the same proportions in the live discard as in the landings. It was therefore assumed that all B2 catch would be of lengths below regulated size limits, and be either age 0 or age 1 in all three subregions during 1982-1996. Catch type B2 was allocated on a semi-annual, subregional basis in the same ratio as the annual age 0 to age 1 proportion observed in the landings during 1982-1996. Mean weights at age were assumed to be the same as in the landings during 1982-1996.

The minimum landed size in federal and most state waters increased to 14.5 in (37 cm) in 1997, to 15.0 in (38 cm) in 1998-1999, and to 15.5 in (39 cm) in 2000. Applying the same logic used to allocate the 1982-1996 recreational released catch to size and age categories during 1997-2000 implied that the recreational fishery released catch included fish of ages 2 and 3. Investigation of data from the CTDEP Volunteer Angler Survey (VAS) for 1997-1999 and from the American Littoral Society (ALS) for 1999, and comparing the length frequency of released fish in these programs with the MRFSS data on the length frequency of landed fish below the minimum size, indicated this assumption was valid for 1997-1999 (MAFMC 2001a). The CTDEP VAS and ALS data, along with data from the NYDEC Party Boat Survey (PBS) was used to validate this assumption for 2000. For 1997-2000 all B2 catch was assumed to be of lengths below regulated size limits, and therefore comprised of ages 0 to 3. Catch type B2 was allocated on a sub-regional basis in the same ratio as the annual age 0 to age 3 proportions observed in the landings at lengths less than 37 cm in 1997, 38 cm in 1998-1999, and 39 cm in 2000.

In 2001, many states adopted different combinations of minimum size and possession limits to meet management requirements. As a result, minimum sizes for summer flounder ranged from 15.5 in (39 cm) in Federal, VA, and NC waters, 16 in (41 cm) in NJ, 16.5 in (42 cm) in MA, 17 in (43 cm) in MD and NY, to 17.5 in (44 cm) in CT, RI, and DE. Examination of data provided by MD sport fishing clubs, the CTDEP VAS, the ALS, and the NYDEC PBS indicated that the assumption that fish released are those smaller than the minimum size remained valid for 2001, and so catch type B2 was characterized by the same proportion at length as the landed catch less than the minimum size in the respective states. The differential minimum size by state has continued since 2001. For 2002-2007, increased samples of the recreational fishery discards by the CT VAS, NYDEC PBS, and the MRFSS For Hire Survey (FHS) has allowed direct characterization the length frequencies of the discards from sample data (Table 85).

Studies conducted to estimate hooking mortality for striped bass and black sea bass suggest a hooking mortality rate of 8% for striped bass (Diodati and Richards 1996) and 5% for black sea bass (Bugley and Shepherd, 1991). Work by the states of Washington and Oregon with Pacific halibut (a potentially much larger flatfish species, but otherwise morphologically similar to summer flounder) found "average hooking mortality...between eight and 24 percent" (IPHC, 1988). An unpublished tagging study by the NYDEC (Weber MS 1984) on survival of released sublegal summer flounder caught by hook-and-line suggested a total, non-fishing mortality rate of 53%, which included hooking plus tagging mortality as well as deaths by natural causes (i.e., predation, disease, senescence). Assuming deaths by natural causes to be about 18%, (an instantaneous rate of 0.20), an annual hooking plus tagging mortality rate of about 35% can be derived from the NYDEC results. In the SARC 25 (NEFSC 1997b) and earlier assessments of summer flounder, a 25% hooking mortality rate was assumed for summer flounder released alive by anglers.

However, two more recent investigations of summer flounder recreational fishery release mortality suggest that a lower release mortality rate is more appropriate. Lucy and Holton (1998) used field trials and tank experiments to investigate the release mortality rate for summer flounder in Virginia, and found rates ranging from 6% (field trials) to 11% (tank experiments). Malchoff and Lucy (1998) used field cages to hold fish angled in New York and Virginia during 1997 and 1998, and found a mean short term mortality rate of

14% across all trials. Given the results of these release mortality studies conducted specifically for summer flounder, a 10% release mortality rate was adopted in the Terceiro (1999) stock assessment and has been retained in all subsequent assessments. Ten percent of the total B2 catch at age is therefore the basis of estimates of summer flounder recreational fishery discard at age. In 2007, the number of fish discarded and assumed dead in the recreational fishery was 60% by number and 25% by weight of the total landed (Tables 82-83, Figure 7).

1.5 Total Catch Composition

NER commercial fishery landings and discards at age, North Carolina winter trawl fishery landings and discards at age, and MRFSS recreational fishery landings and discards at age totals were summed to provide a total fishery catch at age matrix for 1982-2007 (Table 88; Figure 8). The percentage of age-3 and older fish in the total catch in numbers has increased during the last decade from only 4% in 1993 to 68% in 2007. Overall mean lengths and weights at age in the total catch were calculated as weighted means (by number in the catch at age) of the respective mean values at age from the NER commercial (Maine to Virginia), North Carolina commercial, and recreational (Maine to North Carolina) fisheries (Tables 89 and 90; Figure 9). The recreational fishery component of the total summer flounder catch has generally increased since 1995 (Table 91; Figure 10).

2.0 Review methods for using fishery-independent surveys as abundance indices in assessment models.

Descriptions of the fishery independent surveys and their associated indices of recruitment and stock abundance are given below. A total of 51 age-specific indices were initially considered as input for the calibration of the assessment modeling frameworks. However, the final base run configurations for each of the modeling approaches under consideration (ADAPT VPA, ASAP, and SS2) included 39 survey indices at age (see section 4.2 for additional detail and discussion).

NEFSC spring

Long-term trends in summer flounder abundance were derived from a stratified random bottom trawl survey conducted in spring by NEFSC between Cape Hatteras and Nova Scotia since 1968 (Clark 1979). NEFSC spring survey indices suggest that total stock biomass last peaked during 1976-1977. The 2007 index (3.17 kg/tow) represents a high series high before falling by over half to 1.41 kg/tow in 2008 (Table 92, Figure 11). Age composition data from the NEFSC spring surveys indicate a substantial reduction in the number of ages in the stock between 1976-1990 (Table 93, Figure 12). For the period 1976-1981, fish of ages 5-8 were captured regularly in the survey, with the oldest individuals aged 8-10 years. From 1982-1986, fish aged 5 and older were only occasionally observed in the survey, and by 1986, the oldest fish observed in the survey were age 5. In 1990 and 1991, only three age groups were observed in the survey catch, and there was an indication that the 1988 year class was very weak. Since 1991, the survey age composition has expanded significantly. There is strong evidence in the 1998-2002 NEFSC spring surveys of increasing abundance of age-3 and older fish, due to increased survival of the 1994 and subsequent year classes. Mean lengths at age in the NEFSC spring survey are presented in Table 94.

NEFSC Autumn

Summer flounder are frequently caught in the NEFSC autumn survey at stations in inshore strata (< 27 meters = 15 fathoms = 90 feet) and at offshore stations in the 27-55 meter depth zone (15-30 fathoms, 90-180 feet) at about the same bathymetry as in the spring survey (Table 92). Furthermore, the autumn survey catches age-0 summer flounder in abundance, providing an index of summer flounder recruitment (Table 95, Figure 13). NEFSC autumn survey indices suggest improved recruitment since the late 1980s, and an increase in abundance of age-2 and older fish since 1995. The NEFSC autumn surveys indicate that the 1995 year class was the most abundant in recent years, and that subsequent, weaker year classes are experiencing increased survival (Table 95). Mean lengths at age in the NEFSC autumn survey are presented in Table 96.

NEFSC Winter

A new series of NEFSC winter trawl surveys was initiated in February 1992 to provide improved abundance indices for flatfish, including summer flounder. The surveys targeted flatfish when they are concentrated offshore during the winter. A modified 36 Yankee trawl was used that differed from the standard trawl employed during the spring and autumn surveys in that long trawl sweeps (wires) were added before the trawl doors to better herd fish to the mouth of the net, and the large rollers used on the standard gear were absent with only a chain "tickler" and small spacing "cookies" present on the footrope.

The design and conduct of the winter survey (timing, strata sampled, and the use of the modified 36 Yankee trawl gear) resulted in greater catchability of summer flounder compared to the other surveys. Most fish were captured in survey strata 61-76 (27-110 meters; 15-60 fathoms) off the Delmarva and North Carolina coasts. Other concentrations of fish were found in strata 1-12, south of the New York and Rhode Island coasts, in slightly deeper waters. Significant numbers of large summer flounder were often taken along the southern flank of Georges Bank (strata 13-18).

Indices of summer flounder abundance from the winter survey indicate stable stock size during 1992-1995, with catch per tow values ranging from 10.9 in 1995 to 13.6 in 1993 (Tables 93 and 97). For 1996, the winter survey index increased by 290% over 1995, from 10.9 to 31.2 fish per tow. The largest increases in 1996 occurred in the Mid-Atlantic Bight region (offshore strata 61-76), where increases up to an order of magnitude occurred in several strata, with the largest increases in strata 61, 62, and 63 off the northern coast of North Carolina. Most of the increased catch in 1996 consisted of age-1 summer flounder from the 1995 year class. In 1997, the index dropped to 10.3 fish per tow, due to the lower numbers of age-1 (1996 year class) fish caught. From 1998-2003, the Winter trawl survey indices increased; with the 2003 Winter survey number and weight per tow indices being the highest in the time series at 35.62 kg/tow (Tables 93 and 97, Figure 11). The Winter survey index was lower from 2004-2007, and values ranged from 12.9 to 21.0 fish per tow. Similar to the other NEFSC surveys, there is strong evidence in recent winter surveys of increased abundance of age-3 and older fish relative to earlier years in the time series (Table 98). Mean length at age in the NEFSC winter survey are presented in Table 99. The Winter survey series ended in 2007.

Massachusetts DMF

Spring and fall bottom trawl surveys conducted by the Massachusetts Division of Marine Fisheries (MADMF) show a decline in abundance in numbers of summer flounder from high levels in 1986 to record lows in 1990 and 1991 (MADMF fall and spring survey, respectively). In 1994, the MADMF survey indices increased to values last observed during 1982-1986, but then declined substantially in 1995, although the indices remain higher than the levels observed in the late 1980s. Since 1996, both the MADMF spring and fall indices have increased to record high levels (Tables 100 and 101, Figure 14). The MADMF also captures a small number of age-0 summer flounder in a seine survey of estuaries, and these data constitute an index of recruitment (Table 102, Figure 15).

Connecticut DEP

Spring and fall bottom trawl surveys are conducted by the Connecticut Department of Environmental Protection (CTDEP). The CTDEP surveys show a decline in abundance in numbers of summer flounder from high levels around 1986 to record lows in 1989. The CTDEP surveys indicate recovery since 1989, and evidence of increased abundance at ages 2 and older since 1995. The 2003 spring and 2002 autumn indices were the highest in the respective time series; although index values decreased in 2004-2007 (Tables 103 and 104, Figure 16. An index of recruitment from the autumn series is available (Table 82, Figure 13).

Rhode Island DFW

Standardized bottom trawl surveys have been conducted since 1979 during the spring and fall months in Narragansett Bay and state waters of Rhode Island Sound by the Rhode Island Department of Fish and Wildlife (RIDFW). Indices of abundance at age for summer flounder have been developed from the autumn survey data using NEFSC autumn survey age-length keys. Survey indices show that the 1984-1987, 1999, 2000, and 2002 year classes are all strong. The autumn survey reached a time series high in 2003 (Table 105, Figure 14). An abundance index has also been developed from a set of fixed stations sampled monthly during 1990-2007. Age-1 indices from this series indicate that strong year classes recruited to the stock in 1996, 1999, 2000, 2002, and 2003 with age 2+ abundance peaking in 2003 (Table 106). Recruitment indices are available from both the autumn (Figure 15) and monthly fixed station surveys.

New Jersey BMF

The New Jersey Bureau of Marine Fisheries (NJBMF) has conducted a standardized bottom trawl survey since 1988. Indices of abundance for summer flounder incorporate data collected from April through October (Table 107, Figure 17). The NJBMF survey mean number per tow indices and frequency distributions were converted to age using the corresponding annual NEFSC combined spring and fall survey age-length keys. Indices of the 1995 year class at age-0 and at older ages in subsequent years indicate that this cohort is the strongest in the time series. Since 1998, most year classes are at or below average; however, the 2005 year class is above average (Figure 18).

Delaware DFW

The Delaware Division of Fish and Wildlife (DEDFW) has conducted a standardized bottom trawl survey with a 16 foot headrope trawl since 1980, and with a 30 foot headrope trawl since 1991. Recruitment indices (age 0 fish; one index from the Delaware estuary proper for 1980 and later, one from the inland bays for 1986 and later) have been developed from the 16 foot trawl survey data. Indices for age-0 to age-4 and older summer flounder have been compiled from the 30 foot headrope survey. The indices use data collected from June through October (arithmetic mean number per tow), with age 0 summer flounder separated from older fish by visual inspection of the length frequency. The 16 foot headrope survey indices suggest poor recruitment in 1983, 1988, and 1993, improved recruitment in 1994-1995, and above average recruitment in 2000 (Tables 108 and 109, Figure 18). The 30 foot headrope survey indices suggest stable stock sizes over the 1991-2001 time series, with strong recruitment in 1991, 1994, 1995, and 2000. The 2004 index from the 30 foot survey was a time series low, along with lower index values from 2002 onwards, with an increase in 2007 (Table 110, Figure 17). These lower index values presumably reflect decreased availability to the survey, rather than a true decrease in abundance.

Maryland DNR

The Maryland Department of Natural Resources (MDDNR) has conducted a standardized trawl survey in the seaside bays and estuaries around Ocean City, MD since 1972. Samples collected during May to October with a 16 foot bottom trawl have been used to develop a recruitment index for summer flounder for the period 1972-2007. This index suggests that weakest year class in the time series recruited to the stock in 1988 and the strongest in 1972, 1983, 1986, 1994, and 1998. The 2001 and 2007 index values were above average, while the 2002 to 2005 values were the lowest values in the last 10 years (Table 111, Figure 18).

Virginia Institute of Marine Science

The Virginia Institute of Marine Science (VIMS) conducts a juvenile fish survey using trawl gear in Virginia rivers and the mainstem of Chesapeake Bay. The time series for the rivers began in 1979. With the Bay included, the series is available only since 1988, but many more stations are included. Trends in the two time series are very similar. An index of recruitment developed from the rivers only series suggests weak year classes recruited to the stock in 1987 and 2005, with strong year classes recruiting during 1980-1984, and 1990, 1991, and 1994. Recruitment indices since 1990 have been below average (Table 112, Figure 19).

North Carolina DMF

The North Carolina Divisions of Marine Fisheries (NCDMF) has conducted a stratified random trawl survey using two 30 foot headrope nets with 3/4" mesh codend in Pamlico Sound since 1987. An index of recruitment developed from these data suggests weak year classes recruited to the stock in 1988 and 2000, with strong year classes in 1987, 1992, and 1996-1998, 2001, and 2002, and 2005 (Table 113, Figure 19). The survey normally takes place in mid-June, but in 1999 was delayed until mid-July. The 1999 index is therefore inconsistent with the other indices in the time series, and the 1999 value was excluded from the VPA calibration in the SARC 31 assessment (NEFSC 2000).

2.1 Evaluate whether to combine several of the surveys into a composite survey index. If appropriate, implement this approach.

2.1.1 Integration of Survey Indices

For this assessment, a working paper was prepared that examined methods to better integrate trends in abundance provided by survey indices, prior to their use in population model calibration. Past peer reviews of the summer flounder assessment (NEFSC 2005), as well as other Northeast species assessments, have recommended investigation of methods to better integrate trends in abundance provided by survey indices (state and federal), prior to their use in population model calibration. These recommendations stem, in part, from the realization that the abundance indices from some state surveys do not index trends for the entire stock, but merely components or substocks of the whole. While some state survey indices may in fact capture stock-wide trends, the peer-review panel research recommendations suggested that a method to statistically summarize and/or appropriately weight indices which are considered *a priori* to adequately characterize stock-wide trends - to “integrate them” - will provide more reliable and transparent results than if the indices were simply used in their original form in Virtual Population Analysis (VPA) calibration. The complete working group paper is provided in Appendix 2. A summary of the methods and conclusions from this paper is provided below.

A GLM approach was used with research survey data to calculate integrated indices of abundance at age for use in a VPA calibration. Data from a recent NER assessment (NEFSC 2005) for summer flounder were used as an empirical test case. The time series of years for the fishery catch and research survey indices was 1982-2003/2004; the VPA calibration used survey indices at age (0-7+) from three seasonal NEFSC trawl survey series and 12 seasonal state surveys. The analytical approach is analogous to a GLM standardization analysis of commercial fishing vessel catch per unit effort data: with the “year” main effect classification variable serves as the index of abundance, while the "survey" classification variable is analogous to a "vessel" classification variable, each with its' own time series of catch per unit effort that has some relationship to the underlying true abundance of the stock. The mean index of abundance is modeled as a log-linear function of the classification variables, with a log-normal error distribution assumption. The analysis could be expanded by including additional classification variables, such as the sampling gear type or tow duration, temporal variables (e.g., spring/fall; day/night) or environmental variables (e.g., water temperature anomalies). However, such details typically are not available for most assessments, and indices are most often presented as aggregate annual or seasonal indices at age. As configured here, the analysis provided average, or integrated, annual indices of abundance at age.

GLM models were constructed for ages 0, 1, 2, 3, 4, and 5-7+. Main effects were limited to the year of sampling (1982, 1983...2004) and the identity of the survey (NEFSC age 1, NEFSC age 2...NEFSC age 5-7+). The resulting year effect coefficients, corrected for lognormal-transformation bias and re-transformed to the original scale were used as a single index of abundance at age input to the VPA calibration in place of the original survey series. Results indicate that without the inclusion in the GLM model of significant main effects (beyond year of sampling and survey identity) that account for a large proportion of the variance of survey series at age from the simple overall means, use of a GLM to develop integrated indices at age provides no clear advantage over using the original indices as input to the VPA calibration. While the GLM integrated indices provide a useful summarization of mean survey trends, the use of integrated indices as VPA calibration input does not guarantee substantially more accurate or precise results than calibration using the original survey indices. The general linear modeling of integrated indices of abundance did provide a useful summarization of mean survey trends. However, the empirical example for summer flounder shows that the use of integrated indices as input to virtual population analysis calibration does not guarantee substantially more accurate or precise results than using the original survey indices.

2.2 Develop and implement an appropriate statistical method to account for the probability of observing zeros in NEFSC survey tows.

The problem of zeros in tuning indices is that a lognormal error distribution is assumed under many assessment frameworks. Since the logarithm of zero is undefined, these zero tuning indices must be either treated as missing data or else be replaced by a positive value. The issue of handling zero observations in the summer flounder assessment tuning indices is not new and has been addressed in a previous Southern Demersal Working Group (SDWG) working paper used in preparing the 2004 summer flounder assessment (SDWG 2003; beginning on page 8). That work responded to the 2002 SAW 35 (NEFSC 2002) summer flounder assessment Research Recommendation: *Explore the sensitivity of the VPA calibration to the addition of 1 and/or a small constant to values of survey series with "true zeros."* In the 2002 (NEFSC 2002) and 2003 (Terceiro 2003a) summer flounder assessments, the addition of the constant value of 1 was made for five age 0 recruitment indices: the MA DMF Seine, CT DEP fall trawl, RI DFW fall trawl, RI DFW monthly trawl, and DE DFW 16 foot bay trawl survey series (note that the latter series was not included in the final ADAPT VPA tuning configuration). No constant was added to survey series with “zero” observations for other age classes. The choice of the value of 1 as the additive constant was based on recommendations from statistical texts (e.g., Snedecor and Cochran 1967, Sokal and Rohlf 1981) for the ln-transformation of data.

Berry (1987) provides guidance on the objective selection of the appropriate value of the additive constant based on the statistical properties (skewness and kurtosis) of data series to be ln-transformed. Work using the procedures suggested by Berry (1987) with recreational fishery catch rates as indices of abundance indicated that the additive constant of 1 was an appropriate value for those data, typically with values between zero and 50 (Terceiro 2003b). The SDWG (2003) work applied the method suggested by Berry (1987) to summer flounder age 0 surveys with “zero” observations. Of the five age 0 series with “zero” observations, the MA DMF series varies between 0 and 70, while the other four series contained small values that varied between 0 and 1. The 2003 work (SDWG 2003) found that for the MA DMF series, the additive constant of 1 minimized the value of *g*. For the other four series, *g* was minimized by small values of the additive constant ranging from 0.001 to 0.1, with an “average” best additive constant of 0.1. The SDWG (2003) therefore recommended use of the revised, varying (1 or 0.1) additive constants in future assessments, and this revision was made in the 2004-2006 assessment, for age 0 survey series only. No constant was added for survey series of other age classes, pending further research.

The 2006 assessment of summer flounder (Terceiro 2006) was subject to a NMFS Office of Science and Technology (S&T) Peer Review (Methot 2006). Among the recommendations made by the S&T Peer Review panel was the following:

The Panel finds that one immediate modification of the VPA is justifiable and reduces the retrospective pattern in stock size during 2003-2005. The VPA model currently treats survey observations of zero as missing values. An observation of zero for a particular age of fish in a particular survey year does not mean that there are no fish of that age in the stock, only that the number of survey samples was not sufficient to detect any fish of that age. This VPA model, as with most assessment models, tunes to the logarithm of the survey observations so cannot explicitly deal with observations of zero. However, treating these zeroes as missing values can result in a bias because time periods of low abundance are underrepresented in the data input to the assessment model. In the case of summer flounder, the result may be an underestimate of the degree to which the stock has rebuilt since the low levels that occurred around 1990. The committee did not discuss this issue during the Sept 14-15 meeting, so is not prepared to present a definitive solution. An interim approach would use a small value in place of the zeroes. A value equal to one sixth of the smallest observed positive value would be reasonable until a more complete statistical solution can be developed.

As a result, the 2006 summer flounder assessment was revised (Terceiro 2006b). The previous treatment of “zero” observations for age 0 indices was retained (additive constant of 1 for MA DMF seine survey, 0.1 for the CT DEP fall trawl, RI DFW fall trawl, RI DFW monthly trawl, and DE DFW 16 foot bay trawl surveys) and age-1-7+ survey observations of zero were replaced with values equal to one sixth of the smallest observed positive value for those series. Typically, the minimum non-zero value in these series was 0.01, and so the additive constant was 0.001667 (Terceiro 2006b).

Therefore, to more fully understand the implications of this recommendation three working group paper were prepared in support of the current assessment to explore methods to address observed zeros in survey indices and to determine how zeros in the tuning indices should be handles in the current assessment. The complete working group papers are provided in Appendix 2. A summary of the methods and conclusions from these papers are provided below. In addition, the WG examined the findings of the ICES working group ICES working group report on this issue entitled “ Report of the Working Group on Methods of Fish Stock Assessment” ICES WGMG Report 2007 (ICES CM 2007/RMC:04).

The first working paper on this subject conducted two types of simulation analyses. The first was a simple spreadsheet example of how a single artificially generated time series is impacted by different levels of fish detection. This artificial population (which exhibited a decline and increase) included values that were rounded to two, one, and zero decimal places creating observations of zero for 2, 4, and 7 years, respectively. A series of constants was added to the time series ranging from 0.0001 to 10 so that the holes were filled and new catchability coefficients were calculated that minimized the difference between the true population and the observed [modified] survey time series. The differences between observed and predicted values depend strongly on the constant added to the time series. However, the more disturbing result demonstrated is that the addition of a constant value to replace the zeros in a survey time series artificially imposed a pattern that may not match the actual pattern in the population. This is most clearly seen in the round 0 case where seven zeros are filled with the same value even though the true population declines then increases during the seven year period.

The second simulation conducted generated many random sets of data for VPA from a known case, created zeros for some of the indices in some years, and compared different methods for dealing with these zeros, and including treating them as missing values, replacing the zeros with a fixed small value, and the one sixth of the smallest observation rule. The simulated population was loosely based on the summer flounder assessment with the population; exhibiting a population decline and increase, spanned 24 years, consisting of 8 age classes (last age class as a plus group), M=0.2, and variable recruitment and F. One index was generated for each age and the suites of identical (age-specific) indices were given four different treatments: Case 1 - actual values used, Case 2 - replaced with 0.0 and treated as missing, Case 3 - replaced with the arbitrary constant 0.01, Case 4 - replaced with 0.0 then a constant of 1/6 times the smallest non-zero element in the index vector added to all index vector elements including zeros. The median values of F and N at age from the 100 realizations of the VPA model under the four cases were compared with the true values from the simulated population. Due to the convergence properties of VPA, the medians from the 4 cases are essentially identical for years 1982-1994. The most striking feature seen is the poor performance of Case 3 (arbitrary constant of 0.01), with values well below the true values while the estimated population abundances were well above the true values, demonstrating the potential for introducing bias by replacing zeros in tuning index time series with an arbitrary constant. While not as clear, generally the Case 4 (add 1/6 of smallest non-zero element) estimates were more biased than the Case 2 (treat zeros as missing) estimates. The exception to this generality is seen in age 1 results where the VPA formulation had to be modified slightly to estimate only ages 3-8 in the terminal year +1 due to the lack of information for age 2 in the terminal year +1 when the index was zero. For older ages, Case 2 actually outperformed Case 1 (all data used) relative to the truth. It is not clear why this happened and may be an artifact of the bias introduced by the mis-ageing matrix used to generate the catch data. However, even if Case 1 is used as the basis for comparison, instead of the true values, Case 2 performs at least as well as Case 4 for all ages except age 1.

An alternative method to determining the constant to use in place of zeros consists of finding the constant that minimizes a function of the skewness plus kurtosis of the raw data Berry (1987). This approach was not considered appropriate for use with tuning index data because the residuals are assumed to follow a lognormal distribution, not the raw observations.

While the 1/6 of the smallest non-zero approach appears to provide reasonable results in some cases, it is an arbitrary rule. In some situations, 1/5 or 1/7 of the smallest non-zero index value would perform better than 1/6. However, filling zeros with a constant value, no matter how that constant is selected, creates a pattern that may not match reality and has the potential to bias the results. The simulations in this working paper demonstrated show that this approach can produce results further from the truth than treating zeros as missing values. In reality, zeros do have information; but results should be checked to ensure that predicted values are not high when index is zero. These two simulation studies demonstrated problems that can arise when tuning indices with zero values are replaced with arbitrary constants. This practice assumes that the correct magnitude can be chosen to fill the zeros and that it is better to provide the model with information that the index is low rather than treat the data as missing. Results demonstrate that this premise is not always correct. Thus, this working paper recommends the NEFSC treat zero values in tuning indices for VPA as missing values.

The second working paper on this subject included a simple regression example to further examine the consequences of adding 1/6 of the smallest non-zero value in tuning series to all values from that series. A 26 year population time series was simulated, with each value varying uniformly between zero and 50,000 fish. Four time series of values were created either with or without a constant of 1/6 of the smallest non-zero value (+c) in the observed time series: ln(obs), ln(obs+c), ln(pred), and ln(pred+c) where ln(obs) was missing when the observed value was zero. Two slopes were computed, one for ln(obs) vs ln(pred) denoted "missing" and the other for ln(obs+c) vs ln(pred+c) deonted "add c." Since in both cases the only source of error is the lognormal error assumed around the observed values, the expectation is that both lines will have slope equal to one. Random series of populations and observation errors were drawn 10,000 times and the two slopes computed for each realization.

When zero observations were treated as missing, the slope was slightly negatively biased (mean 0.983; 90% CI (0.864, 1.109)) and when a constant of 1/6 the smallest non-zero value was added to all observed and predicted values, the slope was highly positively biased (mean 1.261; 90% CI (1.018, 1.483)). Note that the 90% confidence interval for the "add c" case does not overlap one and has a range nearly twice as large as the "missing" case. Under this regression example, the constant was added to both observed and predicted data because to ensure an appropriate comparison. However, in a separate simulation the author did not replace values less than 0.5 with zero and found nearly identical distributions for the "missing" and "add c" slopes; which demonstrated that filling of zeros causes the problem, not the addition of a constant. In order for the "add c" approach to be unbiased, the constant would have to be selected for each realization such that the average of the ln(pred+c) was the same as ln(obs+c) for the values when obs=0. This cannot happen because the predicted values are positive while the observed values are by definition set to zero. Thus, adding a constant to all values when a zero is in the time series will always bias the results. Therefore, the regression example documented in this working paper suggest that filling observed zeros in tuning indices causes a bias relative to the true population that is much greater than the bias introduced by treating the zeros as missing in this simple regression example.

A third working paper on this subject applied the Berry (1987) approach to the summer flounder survey series for all ages with observed "zeros" to determine the best additive constant to use to remove these "zero" observations from the ADAPT VPA calibration data. There were 24 survey series examined and these are characterized by non-zero values between 0.001 and 70, CVs that generally exceed 100%, positive skewness (long right hand tail), and significant kurtosis (high degree of peak, or contagion, near the mean). The proportion of "zeros" in the time series ranged from 1 of 31 = 3% (NEFSC Spring Age 3 index) to 13 of 28 = 46% (MA Fall 4).

Briefly, the methodology of Berry (1987) consists of 1) addition of a range of constants from very large (e.g., 100) to very small (e.g., 0.0001) to the original values in the series, 2) ln-transformation of the modified series, 3) calculation of the skewness and kurtosis of the modified series, and 4) summation of the absolute value of the skewness and kurtosis (providing the statistic g) of the modified series. The additive constant that minimizes g for a given series of data is the one that best minimizes the effect of outliers and normalizes residuals from the lognormal error distribution, hence best adhering to the assumption of the lognormal distribution. These methods applied to the 24 series produced values of g that were minimized for constants between 0.001 and 100, for the age 0, 1, 2, 3, 4, and 5-7+ (aggregate) survey indices (number per tow or haul). There was no statistically significant correlation between the value of the additive constant that minimizes g and the given statistical parameters. Examination of these results for the age-specific indices demonstrated that there is no consistent pattern in the identification of the additive constant that minimizes the absolute value of Berry's (1987) g statistic. There is no strong relationship between the absolute magnitude of the index values, the length of the time series, the number of zeros, the magnitude of the smallest observed value, or any of the usual statistical moments of the series (mean, maximum, non-zero minimum, CV, skewness, kurtosis), and the value of the additive constant that minimizes g. Further, while the "one-sixth" of the minimum observed value was identified as the "best" additive constant in 5 of the 24 (21%) cases examined, this level is not high enough to justify this approach as a reliable rule-of-thumb. In fact, the additive constant of 0.01 was identified as "best" for a higher percentage of series (6 of 24 = 25%). Given the inability to identify a constant that consistently minimizes *g*, the best rule is to maintain the current approach of making no adjustment and continue to treat "zero" observations as "missing."

The three working group papers developed for the current assessment all suggest that it is more appropriate to treat "zero" observations as "missing" in the survey series. The ICES "Report of the Working Group on Methods of Fish Stock Assessment" ICES WGMG Report 2007 (ICES CM 2007/RMC:04), suggests that it may be more appropriate to change the models to assume an error structure other than lognormal, as opposed to filling zero values in the survey series with small positive values.

3.0 Evaluate the feasibility of implementing alternative approaches to assess status of summer flounder stock and comment on any potential effects on estimates of F, SSB, and BRPs.

BIOLOGICAL DATA

Aging

Work performed for the SAW 22 assessment (NEFSC 1996b) indicated a major expansion in the size range of 1-year old summer flounder collected during the 1995 and 1996 NEFSC winter bottom trawl surveys. This also brought to light differences between ages determined by NEFSC and NCDMF fishery biology staffs; therefore, age structure (scale) exchanges were performed after the SAW 22 assessment to explore these differences. The results of the first two exchanges, reported at SAW 22 (NEFSC 1996b), indicated low levels of agreement between age readers at the NEFSC and NCDMF (31 and 46%). In 1996, research was conducted to determine inter-annular distances and to back-calculate mean length at age from scale samples collected on all NEFSC bottom trawl surveys (winter, spring and fall) for comparison with NCDMF samples. While mean length at age remained relatively constant from year to year, inter-annular distances increased sharply in the samples from the 1995-1996 winter surveys, and increased to a lesser degree in samples from other 1995-1996 surveys. As a result, further exchanges were suspended pending the resolution of an apparent aging problem.

Age samples from the winter 1997 bottom trawl survey, aged utilizing both scales and otoliths by only by one reader, indicated a similar pattern as the previous two winter surveys (i.e., several large age 1 individuals), and some disagreement between scale and otolith ages obtained from the same fish. Because of these problems, a team of five experienced NEFSC readers was formed to re-examine the scales aged from the winter survey. After examining several hundred scales, the team determined that re-aging all samples from 1995-1997 would be appropriate, including all winter, spring, and fall samples from the NEFSC and MA DMF [should this be NCDMF?] bottom trawl surveys and all samples from the commercial fishery. The age determination criteria remained the same as those developed at the 1990 summer flounder workshop (Almeida et al. 1992) and described in the aging manual utilized by NEFSC staff (Dery 1997). Only those fish for which a 100% agreement of all group members was attained were included in the revised database, however. The data from the re-aged database were used in analyses in the SAW 25 assessment (NEFSC 1997b).

A third summer flounder aging workshop was held at the NEFSC in February, 1999, to continue the exchange of age structures and review of aging protocols for summer flounder (Bolz et al. 2000). Participants at this workshop concluded that the majority of aging disagreements in recent NEFSC-NCDMF exchanges arose from the interpretation of marginal scale increments due to highly variable timing of annulus formation, and from the interpretation of first year growth patterns and first annulus selection. The workshop recommended regular samples exchanges between NEFSC and NCDMF, and further analyses of first year growth. Recently, Sipe and Chittenden (2001) concluded that sectioned otoliths were the best structure for aging summer flounder over the age range from 0 to 10 years. Since 2001, both scales and otoliths have routinely been collected in all NEFSC trawl surveys for fish larger than 60 cm, and studies are underway to determine the best structure to use for aging these large summer flounder. An exchange of NEFSC and NCDMF aging structures for summer flounder occurred again in 2006, after the SAW SDWG listed the age sample exchange as a high research priority. This exchange examined samples from fish aged 1 to 9 (23-76 cm total length) and determined that the current consistency of aging between NCDMF and the NEFSC is at an acceptable level.

Maturity

The maturity schedule for summer flounder used in the 1990 SAW (SAW 11) and subsequent stock assessments through 1999 was developed by the SAW 11 Working Group using NEFSC Fall Survey maturity data for 1978-1989 and mean lengths at age from the NEFSC fall survey (G. Shepherd, NEFSC, personal communication, July 1, 1990; NEFSC 1990; Terceiro 1999). The SAW 11 work indicated that the median length at maturity (50^{th} percentile, L_{50}) was 25.7 cm for male summer flounder, 27.6 cm for female summer flounder, and 25.9 cm for the sexes combined. Under the aging convention used in the SAW 11 and subsequent assessments (Smith *et al.* 1981, Almeida *et al.* 1992, Szedlmayer and Able 1992, Bolz *et al.* 2000), the median age of maturity (50^{th} percentile, A_{50}) for summer flounder was determined to be 1.0 years for males and 1.5 years for females. Combined maturities indicated that at peak spawning time in the autumn, that 38% of age-0 fish are mature, 72% of age-1 fish are mature, 90% of age-2 fish are mature, 97% of age-3 fish are mature, 99% of age-4 fish are mature, and 100% of age-5 and older fish are mature. The maturities for age-3 and older were rounded to 100% in the SAW 11 and subsequent assessments.

In the past series of summer flounder assessments, it has been noted that the NEFSC maturity schedules have been based on simple gross morphological examination of the gonads; therefore, they may not accurately reflect (i.e., may overestimate) the true spawning potential of the summer flounder stock (especially for age-0 and age-1 fish). It should also be noted, however, that spawning stock biomass (SSB) estimates based on age-2 and older fish show the same long term trends in SSB as estimates which include age 0 and 1 fish in the spawning stock. A research recommendation that the true spawning contribution of young summer flounder to the SSB be investigated has been included in the resulting research recommendations from summer flounder stock assessments since 1993 (NEFSC 1993). In light of the completion of a URI study to address this research recommendation, the maturity data for summer flounder for 1982-1998 were examined in the 2000 assessment (NEFSC 2000) to determine if changes in the maturity schedule were warranted.

The research at the University of Rhode Island (URI) by Drs. Jennifer Specker and Rebecca Rand Merson (hereafter referred to collectively as the "URI 1999" study) attempted to address the issue of the true contribution of young summer flounder to the spawning stock. The URI 1999 study examined the histological and biochemical characteristics of female summer flounder oocytes (1) to determine if age-0 and age-1 female summer flounder produce viable eggs, and (2) to develop an improved guide for classifying the maturity of summer flounder collected in NEFSC surveys (Specker *et al.* 1999, Merson *et al.* 2000, Merson *et al.* In review). The URI 1999 study examined 333 female summer flounder (321 aged fish) sampled during the NEFSC Winter 1997 Bottom Trawl Survey (February 1997) and 227 female summer flounder (210 aged fish) sampled during the NEFSC Autumn 1997 Bottom Trawl Survey (September 1997) using radioimmunoassays to quantify the biochemical cell components characteristic of mature fish.

The NEFSC and URI 1999 maturity determinations disagreed for 13% of the 531 aged fish, with most (10%) of the disagreement due to NEFSC mature fish classified as immature by the URI 1999 histological and biochemical criteria. The URI 1999 criteria indicated that 15% of the age-0 fish were mature, 82% of the age-1 fish were mature, 97% of the age-2 fish were mature, and 100% of the age 3 and older fish were mature. When the proportions of fish mature at length and age were estimated by probit analysis, median length at maturity (50^{th} percentile, L_{50}) was estimated to be 34.7 cm for female summer flounder, with the following proportions mature at age: age-0: 30%, age-1: 68%, age-2: 92%, age-3: 98%, and age-4: 100%. Median age of maturity (50^{th} percentile, A_{50}) was estimated to be about 0.5 years. Based on this new information, SARC 31 (NEFSC 2000) considered 5 options for the summer flounder maturity schedule for the 2000 stock assessment:

1) No change, use the maturity schedule for combined sexes as in the SAW 11 and subsequent assessments (rounded to 0.38, 0.72, 0.90, 1.00, 1.00, and 1.00 as in the SAW 25 and Terceiro (1999) assessment analyses).
2) Consider only age-2 and older fish of both sexes in the SSB.
3) Knife edged, age-1 and older maturity for both sexes. This would eliminate age-0 fish of both sexes from the SSB, and assume that the proportions mature at age-1 "round" to 100%.
4) NEFSC 1982-1989, 1990-1998 for both sexes, assuming a 1:1 sex ratio in deriving a combined schedule.

5) NEFSC 1982-1989, 1990-1998 for males, URI 1999 for females, assuming a 1:1 sex ratio in deriving a combined schedule.

The 5 options produce the following maturity schedules for both sexes combined:

Option			Age			
	0	1	2	3	4	5+
1	0.38	0.72	0.90	1.00	1.00	1.00
2	0.00	0.00	0.90	1.00	1.00	1.00
3	0.00	1.00	1.00	1.00	1.00	1.00
4	0.45, 0.45	0.88, 0.82	0.97, 0.93	1.00, 0.98	1.00, 0.99	1.00, 1.00
5	0.29, 0.31	0.74, 0.76	0.95, 0.94	0.99, 0.98	1.00, 1.00	1.00, 1.00

SARC 31 concluded that some contribution to spawning from ages 0 and 1 should be included, eliminating options 2 and 3. The differences among remaining options 1, 4, and 5 were considered to be relatively minor, and so the SAW 11 schedule (Option 1) was retained for subsequent assessments (MAFMC 2001a, NEFSC 2002). SARC 31 recommended that more biochemical and histological work should be done for additional years to determine if the results of the URI 1999 study will be applicable over the full VPA time series. SARC 31 also noted the need for research to explore whether the viability of eggs produced by young, first time spawning summer flounder is comparable to the viability of eggs produced by older, repeat spawning summer flounder. In the 2005 SAW 41 work (NEFSC 2005), the maturity schedule was updated and broadened to include data from 1992-2004, covering the year range for individually measured and weighed fish sampled in NEFSC research surveys. The resulting combined sex maturity schedule (0.38, 0.91, 0.98, 1.00, 1.00, and 1.00; respectively for age-0 to 5+) was retained in the 2006 assessment and S&T peer review (Terceiro, 2006b).

The SDWG examined the proportions of summer flounder mature at age from 1981-2007 as well as information on length and age at maturity from 1992-2007, and concluded that it was appropriate to retain the maturity schedule from the 2006 assessment. Using NEFSC Fall Survey maturity data from 1992-1997 and probit analysis, the median length at maturity (50^{th} percentile, L_{50}) was estimated as 27.0 cm for male summer flounder, 30.3 cm for female summer flounder, and 27.6 cm for the sexes combined. The median age of maturity (50^{th} percentile, A_{50}) for summer flounder was determined to be 1.1 years for males, 1.4 years for females, and 1.2 years for both sexes combined. These findings are consistent with the findings of SAW 11 and the URI 1999 study. In addition, an examination of the proportions of mature age-0 and age-1 fish did not indicate any trend which would warrant modification of the current maturity schedule (Figures 20 and 21).

3.1 Alternative approaches could consider separate Catch at age matrices for commercial and recreational fisheries, and resulting partial recruitment vectors for each fishery.

Alternative approaches could consider separate catch at age matrices for commercial and recreational fisheries, and resulting partial recruitment vectors for each fishery. The SDWG considered the use of a single catch at age matrix as has been done in previous assessments, two matrices for retained and discarded components of the fishery, and as many as six matrices. These considerations are described in Section 5.0 of this report.

3.2 Alternative approaches could consider regional differences (north, south) in catch at age matrices.

Exploratory analyses were conducted to examine if indeed some of the patterns observed in the summer flounder assessment (Terceriro 2006a, 2006b; i.e. retrospective pattern, large positive residuals primarily at ages 3 and 4 in NEFSC winter, spring, and CT, RI, and NJ indices) could be explained by changes in the spatial distribution of the commercial fishery and or the summer flounder stock. Therefore, the commercial landings data from 1967-2006 and NEFSC survey data were examined spatially. It should be noted that this data was compiled from a generalized data retrieval, therefore this numbers may differ slightly (within a few tons) from the assessment data tables.

Commercial landings were compiled by year, regional “division”, and calendar quarter (1-4). While the data were examined by calendar quarter, for simplicity, the discussion of here focuses only on year and “division”. These “divisions” were aggregations of the NEFSC commercial fishery statistical areas (SAs; Figure 22; units generally about 1 degree square) to allow for better investigation of regional differences. The following logical divisions were created based on the aggregation of statistical areas; Division 51 (Gulf of Maine) aggregates SAs 511-515; Division 52 (George’s Bank) aggregates SAs 521-526, 561, and 562; Division 53 (Southern New England) aggregates SAs 533-539; Division 61 (Northern Mid-Atlantic Bight) aggregates SAs 611-616; Division 62 (Southern Mid-Atlantic Bight) aggregates SAs 621-629; Division 63 (Virginia/North Carolina region) aggregates SAs 631-636; and Division 99 is all other landings outside these 6 regions (SAs not reported).

There are three time periods worth noting that may influence the pattern exhibited in the commercial landings data: 1967-1981; 1982-1992, and 1993-2006. The period 1967-1981 is prior to any collection of the commercial fishery length-age-composition data, and is the period before the first comprehensive management measures were enacted in 1982. In addition, for 1967-1981 the set of commercial landings with SAs available is incomplete, relative to the total commercial landings reported, due to limited participation of the states in the Federal data collection program (i.e. the "weighout" system). As shown in Table 114, Division totals for the 1960s and 1970s are well below the commercial total reported in the assessment tables, and are dominated by landings report for Divisions 52, 53, and 61. Most states did not fully participate until 1982; New York and New Jersey began participating in 1986 and North Carolina joined in 1997. Therefore, comparing this earliest period (pre-1982) with more recent two periods (post-1982) is not consistent.

The time period 1982-1992 represents an era of pre-quota management; quota management (output controls) have only been in place since 1993. These regulator changes may be reflected in changes in the Divisions accounting for most of the commercial landings between these periods, if substantial changes in the spatial distribution of the commercial fishery were occurring.

In addition, these three periods roughly coincide with the pre-age-structured assessment period (1967-1981), the period of time when the stock is estimated to have been decline (1982-1992), and the period during which rebuilding plans have been in place and the stock has been expanding (1993-present). The overall expansion, contraction, and expansion again of the length and age structure of the stock would be expected to coincide with these three periods. Given the migratory nature of summer flounder, and of the behavior of larger/older fish migrating further North, one might expect to see some changes in the spatial distribution of large fish (e.g. ages 3 and 4, and older) in both the commercial data and survey over these periods (survey follows later). The divisions that accounted for the greatest proportion of commercial landings in a given year during the latter two time periods (1982-1992; 1993-2006) were Division 61 and 62, the areas of the Northern and Southern Mid-Atlantic Bight, respectively. Division 61 accounted for the highest commercial landings 9% of the time (1 out of 11 years) for the period 1982-1992, while Division 61 accounted for the highest landings 29% of the time (4 out of 14 years) during the period 1993-2006. This suggests a shift in the fishery northwards (Figure 23). Drawing further insight from this exploratory work would require de-construction of the commercial and recreational expanded length and age frequencies; one would expect to find larger/older fish in the commercial landings from the more Northerly Divisions (mostly Division 53–Southern New England), compared to the more Southerly Divisions (61-63). There may however, be limits to this exercise depending on the sample sizes for the length-age data during some years of the time series, particularly the earlier and mid-parts of the data collection series and if the additional factor of calendar quarter is included. In addition to examining the commercial landings data, the NEFSC survey data may provide insight into some of the patterns observed in the summer flounder assessment.

The NEFSC Spring survey data was compiled into two regions. These “regions” were aggregations of the NEFSC Spring survey strata (strata; Figure 24), with a Northern region aggregating strata 1-12 and a Southern region aggregating strata 61-76. These offshore strata are the standard suite of those used for calculation of that spring tuning index. Because of the low numbers of summer flounder caught in the strata on George’s Bank, those strata are not used in calculation of the tuning indices and are not used in the exploratory analysis presenting here (Figure 25). As shown by Figure 26, indices of abundance over most of the last 40 years have generally been higher in the South, while the indices of biomass are comparable for the Northern and Southern strata. This is consistent with the expectation of older/larger fish being found in the North. Since around 2000, with the exception of 2004, both abundance and biomass have been higher in the

North. An examination of length-frequencies for the period 2002-2004 by region (Northern versus Southern strata; Figure 27) are consistent with these findings and suggest that most of the differences in length frequencies appear between 40-50 cm, which are age-3 and age-4 fish.

The survey data do provide evidence of more older/larger fish being found in the NEFSC spring survey “Northern” strata (1-12) since around 2000. In addition, many of the Northern state-specific surveys show a similar pattern. This pattern appears to be more evident in the survey data than was shown in the commercial landings data examined. This exploratory analysis suggests that the development of assessment models which include a regional (spatial) component may be worthwhile and should continue to be included as a future research recommendation. It should be noted that this recommendation may be in conflict with TOR 2.1, which suggests development of integrated survey indices for use in the assessment.

3.3 Alternative approaches could consider potential gender differences in life span, growth rate, and natural mortality and implications of these factors for observed age- and length-specific sex ratios.

During the 2007 stock assessment update (SDWG 2007), it was noted that there is potential for change in certain biological parameters of the stock over the last few years. Therefore, a working group paper were developed (and summarized below) for this benchmark assessment to examine a variety of biological parameters and if there are changes in length-at-age, weight-at-age, and growth rates, and sex ratios (Working Papers 7 and 9; Appendix 3).

The first working paper examined trends in the NEFSC trawl data. Catches from the NEFSC trawl survey database (1992-2007) are subsampled and provide length, weight, age, and sex for summer flounder. Sample size at older ages was low, particularly during early years of this analysis corresponding to periods of lower abundance. The data were therefore limited to years and ages that had a sufficient sample size. Length at age calculations were developed from the NEFSC winter survey only and include 1999 to 2006, ages 0 through 4 for males, and 0 through 5 for females. Sample size for these years and ages are generally greater than 40 fish. Sex specific mean size at age was calculated for each year and SAS Proc REG (SAS 1990) was used to conduct regressions of size at age over time. The observed data were fit to a von Bertalanffy growth function using SAS Proc NLIN (SAS 1989a). Residuals were then resampled with replacement, by year, to develop 500 bootstrap datasets (Barker 2005), each of which was also fit using the von Bertalanffy growth function. Similarly, length-weight analysis was conducted using an allometric growth function and potential changes in weight-at-age were examined.

Mean lengths for males age 1 to 4 show no trends (given the limited data 1999-2006), and regression trends were not significant ($\alpha = 0.10$ level). Trends in mean length at age for females were similar to males for ages 0 to 4; however, female mean length at age 5 decreased significantly between 1999 and 2006.Fitting bootstrap data to the von

Bertalanffy growth function resulted in unrealistic parameter estimates for males in 2000 and 2006 ($L\infty$=100,000 cm) and for females in 2000. Regression results, using the von Bertalanffy estimates, for male length at ages 0 to 10 (the approximate age range observed in survey data) showed no significant trends. Regression results indicate no significant trend in predicted length at age for females ages 0 to 4; however, predicted length at ages five and older decreased significantly between 1999 and 2006. Combined sex regression results using von Bertalanffy predicted length at age were consistent with results of mean length at age. Length:weight analysis was conducted using the same subset of years and ages described above. Sample size was generally above 40 fish and no significant trends were observed for weight at length for males or females.

Maximum age, as identified through review of NEFSC spring, winter, and fall survey data, indicated that the maximum age for males generally varied between age 4 and 5 from 1985-1995, while female maximum age ranged from age 6 to 8. By 2000, the maximum age of males increased to between 8 and 9, where it remained stable until 2007 when one 12 year old male was captured. Female maximum age has increased steadily since 1995, with a peak of 14 years in 2005. Dery (1988) suggested males and females reached maximum ages of 7 and 12 years, respectively. While this is consistent with maximum ages observed in NEFSC trawl surveys from 1992 to 2000, recent data suggest that maximum ages of at least 12 for males and 15 for females is more appropriate. Additional years of reduced fishing pressure may result in even older observed maximum ages.

Sex ratio (*i.e.* percent female) at size was analyzed using SAS Proc GENMOD (SAS 1989b) with a normal distribution and a logit link function (*i.e.* a logistic regression). From 1992 to 1997 overall sex ratio was about 54% female, then increased from 53 to 58% of the stock in 1997 to 2000, where it remained stable for 3 years. In 2003, the ratio dropped to 51% female and has varied in recent years. Sex ratios by age showed a decrease in percent female since the mid 1990s across all ages, although the declines are more evident for ages 2+ .When data are combined across years, logistic regression of percent female at size shows a 50:50 sex ratio at around 38 cm (15”). Fish smaller than this size are predominantly male, while larger fish are predominantly female. These findings are not new (*e.g.* Murawski and Figley 1977, Morse 1981), but this sexual dimorphism may have greater implications for management and stock rebuilding. For many ages, this decline has been observed over 15 years which is much longer than states have required large minimum sizes.

Natural mortality for each sex was estimated using $3 / T_{MAX}$ (Hoenig 1983), assuming maximum ages of 12 and 15 for males and females, respectively, and resulted in M=0.25 for males and M=0.20 for females Applying these to overall sex ratios to estimate annual M, M has remained relatively stable around 0.223, with a range of 0.221 in 2000 to 0.226 in 2005. A more comprehensive examination of methods to estimate M is available in Section 3.4.

Individual fecundity was estimated by applying the relationship of Morse (1981) to the mid-year length at age for females.

$$F = 0.0007975 * L^{3.402}$$

Mature females by age and year (in numbers) were determined by multiplying the VPA estimated abundance, sex ratios, and VPA input maturity schedule. Fecundity could only be evaluated for the years 1999 to 2006 due to low samples sizes. Theoretical fecundity of the stock increased from approximately 22.3 x 10^{12} eggs in 1999 to a maximum of over 36.5 x 10^{12} eggs in 2004, and decreased in to approximately 31.0 x 10^{12} eggs in 2006. Recruitment as calculated in the VPA remained relatively stable between 28 million and 38 million individuals, except for 2004 (17 million). The relationship between fecundity and recruitment is slightly negative, although this appears to be driven primarily by the 2004 data point (highest fecundity and lowest recruitment). It appears the increases in total abundance have outpaced any decreases in fecundity, resulting in theoretical stock fecundity increasing more than 50% from 1999 to 2004. Estimated fecundity declined in 2005 and 2006, coincident with slower stock growth; however, it is not clear there is a causal relationship.

In conclusion, this review of NEFSC trawl survey data do indicate that some life history parameters have changed since 1992; although many of the causal relationships have not be established. This descriptive information can be considered by the SDWG in development of the assessment, reference point calculations, and model projections.

A second working paper (summarized below) was developed to evaluate to describe information in the summer flounder biological data base (Working Paper 9; Appendix 3). This work attempted to answer the questions of whether the current data support development of and use of: 1) a sexually-explicit model for summer flounder 2) regionally-specific sex-at-age keys 3) differential natural mortality rate for male and female summer flounder or a nonlinear whole-stock natural mortality rate, or 4) regionally-specific age-length keys.

At the time of analysis, sex ratio data for young-of-the-year are not available prior to 1982; consequently analyses of sex ratio focus on 1982-2007. Due to data limitations, and regional variability in sex ratios as discussed in a subsequent section, data were excluded from southern New England north and also from Cape Hatteras south in this set of analyses as well as all age-year combinations where the number of sexed summer flounder is less than 20. For some analyses, data were parsed into six year groups with the central four being half-decadal (i.e. Year group 1 contains data from 1982-1985). The data suggest that young-of-the-year are dominantly male. A female-biased sex ratio for young-of-the-year summer flounder occurs only thrice in 26 years. The data also suggest a consistent change in sex ratio with age. Thus, summer flounder are consistently characterized by biased sex ratios regardless of age or half-decadal period within the time series and the direction of bias changes with age. In addition, the three years where females predominate in age-0 fish include the last two years. This is unexpected from the time series record. However, the sex ratio for age-1 fish from the 2006 cohort conforms with age-1 sex ratios from other years; thus, it is possible the 2006 young-of-the-year ratio is a sampling artifact. Second, the fraction of fish that are male at older age has increased over time, although remaining well below 0.5. This is particularly apparent for age-3 fish.

One explanation is that male fish are moderately more susceptible to the fishery at high fishing mortality rates. The dispersion of males and females as the cohort ages might support this first alternative. The same outcome would be obtained either if a reduction in natural mortality rate had occurred if the originating sex ratio was biased to a greater degree in favor of males. A number of potential reasons exist for the male-dominated sex ratios seen in young-of-the-year summer flounder. Females mature later than males. The observed females may under-represent the total number. The biological database also records undifferentiated fish. Assigning all of these fish to the female sex, however, does not markedly change the sex ratios summarized in the data. This suggests the maturity schedule along cannot explain the male-dominated sex ratios observed for age-0 fish. Young males may be more available to the survey. While this possibility cannot be excluded, the fact that females grow faster than males and that the male-biased sex ratios clearly are retained into age 2, albeit at diminishing intensity, suggest that availability is not an adequate explanation. Protandry would produce the observed age-dependent sequence of sex ratios. Protandry, however, is not reported in flatfish, and would almost assuredly have been observed were it to exist. Biased sex ratios have been observed by others in summer flounder, however. Morse (1981) and Smith and Daiber (1977) found that younger, smaller fish were much more likely to be male and that this trend quickly reversed with increasing age. Morse (1981) offers that an initially male-dominated sex ratio is necessary to offset an apparently higher natural mortality rate in males, thus promoting a more nearly 1:1 sex ratio in the spawning stock. One viable explanation for biased sex ratios in young-of-the-year summer flounder is temperature-dependent sex determination; this is consistent with some findings in the literature and has been observed in other flatfishes.

To examine this further by age, depth, region, and time, sex-ratio strata were allocated to three depth zones (<25 fm; 25-50 fm; >50 fm), five regions: southern New England (we included Georges Bank strata in this grouping), the northern Mid-Atlantic Bight, Delmarva, and the strata south of Cape Hatteras, and allocated to half-decadal year-groups: 1976-1980, 1981-1985, 1986-1990, 1991-1995, 1996-2000, 2001-2005, 2006-2007. We excluded all occurrences of age-year group, age-region, and age-depth combinations with sex ratios supported by a total count of males and females less than 30. ANOVAs were run by age using depth, year-group, and region as main effects. All interaction terms were included. Sex was implemented as a dependent variable by assigning a 0 to males and a 1 to females. Means, accordingly, were equivalent to the fraction female.

Examining the regional results for overall trends, it seems that the northern Mid-Atlantic and Delmarva regions have similar sex ratios regardless of age. In addition, the south Atlantic and southern New England regions have a tendency to be different from the Mid-Atlantic/Delmarva grouping, depending on age. When different, the southern New England and south Atlantic regions routinely have a higher fraction of males. This is what would be expected from the temperature-dependent determination of sex that produces an increase in fraction male at the temperature extremes; however, sex determination in the first year of life militates against this explanation as the main effect of region is observed only later. Thus, alternative biological explanations or determinants from differential fishing mortality must be sought. The depth and year-group effects are, as yet, unexplained.

The gradual shift in sex ratio from male-dominated to female-dominated with increasing age might accrue from differential mortality or differential availability. A higher natural mortality rate in males could potentially be explained by some type of biological refuge for females; less time spent in a predatory window. Female summer flounder are known to grow at a faster rate than males and may therefore be less prone to predation (Poole 1961). However, male and female growth rates are similar until age 2, so such an explanation would not be warranted when considering the apparent differential natural mortality in fish younger than age 2. Some precedent exists for higher morality rates in male relative to female flatfish. Morse (1981) already proposed a higher natural mortality rate for males in summer flounder. Santos (1994) computed natural mortality rates for the four-spot megrim (*Lepidorhombus boscii*) by sex. The natural mortality rate for males was 0.41, and for females 0.34. Pearson and McNally (2005) also calculated mortality rates, using three different methods, for the sand sole, *Psettyichthys melanostictus*. The natural mortality rate for females ranged between 0.35 and 0.45, whereas the mortality rate for males was estimated to fall between 0.40 to 0.60.

Additional non-paramentric analyses (categorical ANOVAs) examined region, depth, and year group, with length as the dependent, ranked variable. Depth significantly impacted length-at-age for males and females, ages 1 through 3 and age 4 for females. At age 0, summer flounder are only present in shallow waters, and at age 5 and older, depth no longer influences length-at-age, for the most part. Tukey's studentized range tests show that fish in deeper water are larger at a given age than fish in shallower water. Length at age varied significantly with region for male and female summer flounder, ages 0-4, but not at older ages. When regions did group together they did so in a north-central, south-central trend. In other words, the southern New England region never grouped with the south Atlantic or Delmarva regions, and the south Atlantic never grouped with the northern Mid-Atlantic and southern New England regions. Year-group consistently affected length at age for male and female summer flounder until age 6. Year-group no longer impacted length at age for male summer flounder at age 6 and older, but continued to do so for female fish.

Not considering the influence of fishing mortality on age-at-length, these trends could suggest that summer flounder either grow at faster rates in deeper water and northern latitudes or that larger fish at age preferentially aggregate in these regions. Alternatively, in shallow waters and at southern latitudes larger fish may be more accessible to the fishery. While the fishery may not keep younger fish due to minimum size restrictions, younger fish may still be removed by the fishery as discard mortality. Whether it be a biological reason (e.g., differential growth rates) or a fishery-related reason (bigger fish at any age are more accessible in shallow/southern water), it seems clear that the average size of fish at age is larger in deeper/northern water than in shallow/southern water. However, significant interaction terms also occur commonly in fish 4 years or less in age

and these involve both depth and region with relatively equal frequency and intensity. The frequency of significant interaction terms including depth and region suggests that regionality in the trends in age at length for summer flounder cannot facilely be explained simply in terms of depth and latitude. A more complex mixture of biology and, perhaps, relative fishing impact is likely to be required.
Analyses were conducted that focused on the age-length keys for the Delmarva and northern Mid-Atlantic region and, independently, on the three depth zones previously described, as these two regions were most similar in length at age. To compare keys efficiently, lengths were combined into 12 units, the central 10 being 5 cm; this yielded three keys for the three depth zones and two for the two regions. Each of these returned a significant result from a by-region or by-depth chi-square test, and from a Cochran-Mantel-Haenszel test controlling for depth or region. To directly compare two keys, we used Geary's C and Moran's I statistics on the set of residuals obtained by calculating the expected key structure in one array from the observed key structure in the other.

From this work, it was concluded that young-of-the-year summer flounder are dominantly male. Sex ratio changes gradually with age such that male frequencies over 0.5 occur infrequently by age 2 and rarely exceed 0.3 by age 4. The biased sex ratio at birth may be the result of temperature-dependent sex determination. In additional, this data suggests the need to implement a sex-explicit and/ or spatially-explicit model for summer flounder. The change in sex ratio with age also suggests that separate natural mortality rates be considered for male and female summer flounder stock assessment models with a higher natural mortality rate in males. Spatial variation in length-at-age suggests that a single age-length key is not likely to be representative across all regions and in different depths. The differential with region and depth suggests that differential fishing pressure cannot be excluded as the mechanism generating these differences. The male age-length relationships are move variable over depth and region than the female ones, but each varies significantly. The analyses suggest that a single age-length key may not adequately describe the stock, particularly for the males.

The SDWG considered the information and concluded that while there were some significant interactions among sex-ratios and length-at-age by depth, region, and time period, the patterns and frequency of significant terms suggest the explanatory variables are not being adequately characterized (grouped). Therefore, additionally work is need to address what factors may be appropriate to characterize the observed patterns before a sex-specific or spatially-explicit model can be developed. There are many factors that may be interacting to cause spatial differences in summer flounder length-at-age and sex-ratios, which could include sampling effects, fishing mortality patterns (fishery behavior within year and over time), temperature, and predation patterns. While these analyses focused on NEFSC survey strata, the patterns observed in the fishery catch data (for which there is no sex-information collect), could be different.

This SDWG did, however, consider the information on natural mortality rates, which is discussed in greater detail in the next section (section 3.4).

3.4 Alternative approaches could consider the strength of evidence for natural mortality rate used in the assessment; Update the estimate if appropriate.

Natural Mortality Rate

The instantaneous natural mortality rate (M) for summer flounder was assumed to be 0.2 in past assessments (SAW 20; NEFSC 1996a). In the SAW 20 work, estimates of M were derived using methods described by: 1) Pauly (1980) using growth parameters derived from NC-DMF age-length data and a mean annual bottom temperature (17.5°C) from NC coastal waters; 2) Hoenig (1983) using a maximum age for summer flounder of 15 years; and 3) consideration of age structure expected in unexploited populations (5% rule, 3/M rule, e.g., Anthony 1982). SAW 20 (NEFSC 1996a) concluded that M = 0.2 was a reasonable value given the mean (0.23) and range (0.15-0.28) obtained from the various analyses, and this value for M has been used in all subsequent assessments.

For this assessment, a working paper was prepared that reviewed longevity- and life-history based estimators of M. These sex and age-specific estimates of M were calculated from current summer flounder age and growth data (1976-2007) from the NEFSC trawl surveys. The complete working group paper (Working Paper 8) is provided in Appendix 3. A summary of the methods and conclusions from this paper are provided here. Longevity based estimators of M are sensitive to critical underlying assumptions which include the value of p, or the small proportion of the population surviving to a given maximum age, and the maximum observed age under no or low exploitation conditions. Using a t_{MAX} (maximum age) of 15 years for summer flounder, and the methods of Hoenig (1983) and Hewitt and Hoenig (2005), longevity based estimates of M for combined sexes ranged from 0.20 to 0.28 depending on whether a p=1.5% or p=5% was assumed. Other life-history based were examined and included Pauly (1980), Jensen (1996), Gunderson & Dygert (1988), and Gunderson (1997), with resulting estimates ranging from 0.20 to 0.45; although again these estimates are highly dependent on their underlying assumptions. Age-specific and size variable estimates of M, based on the work of Peterson & Wroblewski (1984), Chen & Watanabe (1989), Lorenzen (1996), Lorenzen (2000), ranged from 0.19 to 0.90, with the highest values obviously associated with age-0-1 fish (fish at smaller lengths). While these exercise provided a wide range of methods and M estimates to be considered, each estimate involved a suite of underlying assumptions which were debated. In addition, the modeling frameworks of SS2 and ASAP (see section 4.2 and 4.3) allow for log-likelihood profiling of M to determine which M estimate provides the best model fit. Based on this exercise using the base cases, M would be 0.25 and 0.20 under the models SS2 and ASAP, respectively.

The SDWG considered the different methods of estimating M and after lengthy discussion assumed a natural mortality rate (M) of 0.20 for females and 0.30 for males for this June 2008 assessment based mainly on recently observed maximum ages (tmax) in NEFSC survey data of 14 years (76 cm, in NEFSC Winter Survey 2005) for females and 12 years (63 cm, in NEFSC Spring Survey 2007) for males, and the expectation that larger and older fish are likely if fishing mortality rates are maintained near current rates over the next several years. The assumptions were aguided by the work above as well as the SDWG working papers on summer flounder growth and maturity prepared by Brust, Powell, and Wong (Working papers 8,9,10; Appendix 3). A combined sex M-schedule at age was developed by assuming these initial M rates by sex, an initial proportion of females at age 0 of 0.40 derived from the NEFSC Fall survey indices by age and sex, and population abundance decline over time at the sex specific M rates. The final abundance weighted combined sex M-schedule at age ranged from 0.26 at age 0 to 0.24 at age 7+, with a mean of 0.25. Additional details on the sensitivity of the assessment to an increase in M are discussed in Section 6.6.

4.0 Compare results from alternative modeling approaches with those from the VPA model, to evaluate the robustness of VPA model results. Perform retrospective analyses of F, SSB, and recruitment for the models, and describe potential effects of retrospective patterns on assessment and rebuilding.

4.1 A Stock Production Model Incorporating Covariates (ASPIC)

The SDWG did not repeat an ASPIC analysis in this assessment. Past attempts to apply this modeling approach to the summer flounder assessment are described in greater detail below; estimates from the model were not considered to be robust and the associated biological reference points were therefore considered to be unreliable. In addition, approaches suggested in a submitted working paper (Working Paper 10, Appendix 4) were not considered by the SDWG to be the most appropriate approach for the current benchmark assessment. The SDWG determined that the extensive age-information available for summer flounder should be utilized for this assessment and other modeling frameworks provide greater flexibility in developing the underlying assumptions, that are implicitly determined based on which surplus production approach is selected.

The non-equilibrium surplus production model incorporating covariates (ASPIC; Prager 1994, 1995) can be used to estimate maximum sustainable yield (MSY) and other biological reference points. An ASPIC analysis applied to summer flounder using various state and federal agency survey biomass indices (the 1998 analysis) was previously reviewed by the NEFMC Overfishing Review Panel (Applegate et al. 1998). Based on total weighted mean squared error (MSE), the NEFSC spring and autumn biomass indices gave the best fit to the data in that analysis. However, the Overfishing Review Panel concluded that biological reference points estimated in the 1998 analysis for summer flounder were unreliable, due to the short time series of reliable catch estimates and lack of dynamic range in the input data (Applegate et al. 1998).

An ASPIC analysis using projected catch and NEFSC survey biomass indices through 1999 was reviewed in the 1999 assessment (Terceiro 1999). Model results were examined for sensitivity by employing a Monte Carlo search routine and by initializing over a broad range the values of MSY (10,000 to 50,000 mt) and the intrinsic rate of increase (r : 0.12 to 1.25). The ratio of initial to current biomass (B1 ratio) was assigned a starting value of 0.50. Overall, the 1999 ASPIC model results for summer flounder were not well defined and suggested the possibility of numerous local minima in the sums of squared errors (SSE) response surface. The Monte Carlo search algorithm was employed in an attempt to provide a better search of the SSE response surface, and the this generated a range of estimates of MSY from 19,000 mt to 58,000 mt and of r from 0.49 to 1.08. Due to the number of iterations needed to reach convergence (>25) and the probable number of local minima, these results also appeared to be unreliable. Thus, biological reference points for summer flounder estimated by the 1999 ASPIC analysis were not considered to be robust.

4.2 ADAPT Virtual population analysis (VPA)

Fishing mortality rates in 2006 and stock sizes in 2007 were estimated using the ADAPT method for calibration of the VPA (Parrack 1986, Gavaris 1988, Conser and Powers 1990) as implemented in the NOAA Fisheries Toolbox (NFT) ADAPT VPA version 2.7.7. As recommended by the MAFMC SSC Committee during the review of the Terceiro (1999) assessment and by the National Research Council review of the summer flounder assessment (NRC 2000), ages 0-6 were included in the analysis as true ages, with ages 7 and older combined as a plus group. An instantaneous natural mortality rate of M = 0.2 was assumed for all ages in all years. Maturities were retained from the last revisions made in the 2005 SAW 41 assessment (NEFSC 2005); maturities at age for all years were 38% for age-0, 91% for age-1, 98% for age-2, and 100% for ages 3 and older. Stock sizes in 2007 were directly estimated for ages 1-6, while the age 7+ group was calculated from Fs estimated in 2006. Fishing mortality on the oldest true age (6) in the years prior to the terminal year was estimated from back-calculated stock sizes for ages 3-6. Fishing mortality on the age 7+ group was assumed equal to the fishing mortality for age 6. Winter, spring, and mid-year (e.g., RIDFW monthly fixed station, DEDFW and NJBMF) survey indices and all survey recruitment (age-0) indices were compared to population numbers of the same age at the beginning of the same year. Fall survey indices were compared to population numbers one year older at the beginning of the next year. Tuning indices were *a priori* unweighted.

A number of ADAPT VPA runs using were made to examine the sensitivity of the analysis to several revisions to data and analyses that have been incorporated in this assessment. These changes include a) revisions to the historical time series of recreational fishery data, including state of North Carolina catch type B2 estimates (run INIT), b) use of the exact catch equation, as opposed to the Pope's approximation used previously (run EXACT), c) revisions to the MADMF trawl survey indices due to changes in stratum area specification (run MASV), d) treatment of survey zero values as missing data (run NOFILL), e) use of only NEFSC survey data in calibration (run NOFILL_NEC), and f) use of all survey indices (run F08_ALL). Of these changes from the 2007 final run configuration, the analysis was most sensitive to the treatment of zeros as missing (change from the MASV to NOFILL run).

A total of 51 age-specific indices were initially considered as input for the ADAPT VPA calibration and other models. The ADAPT VPA was used as the platform to select the base set of indices to carry forward because the existing NFT ADAPT software has very useful diagnostic features for judging the calibration performance of the indices. The inclusion of each survey index was considered based on a pre-calibration correlation analysis among all indices, a post-F08_ALL indices run correlation analysis among the indices and resulting ADAPT VPA estimates of stock size, and an examination of the analytical diagnostics (including the precision of each survey index series, the partial variance accounted for by each index, patterns in residuals, and the mean squared residual (MSR) of the calibrated solution). Survey indices with trends that did not reasonably match corresponding patterns in abundance as estimated by other indices and/or the F08_ALL run were eliminated from the tuning configuration.

The DEDFW 30 foot trawl indices of abundance were considered to be reflective of local population trends and not stock level trends. On that basis and the large amount of variance that the DEDFW indices contributed to the overall model fit, the DEDFW 30 foot trawl ages 0-3 indices were dropped. The MADMF spring age 1, MADMF fall age 1 (tuned to age 2), CTDEP spring age 1, and RIDFW fall age-1 (tuned to age-2) indices correlated poorly with other regional indices and exhibited large CVs and partial variances and were also dropped. The RIDFW fall, MADMF seine survey, DEDFW 16 foot Estuary trawl, and NCDMF age 0 (young-of-year) indices were also dropped due to large partial variance and poor correlation with F08_ALL estimates of recruitment.

The final base run configuration (run F08_BASE) includes 39 survey indices at age. This base set of calibration indices was also used in subsequent runs of the other age-structured models (ASAP and SS2) considered in the assessment. Figures 28-30 compare the estimates of Spawning Stock Biomass (SSB; mt), recruitment at age 0 (R; 000s) and fully-recruited fishing mortality rate (F, ages 3-5) from the alternative ADAPT VPA run configurations with the final F08_BASE run.

The annual partial recruitment of age-1 fish decreased from near 0.50 during the first half of the VPA time series to less than 0.30 since 1994, and to less than 0.20 during 2000-2006; the partial recruitment of age-2 fish has decreased from 1.00 in 1993 to about 0.50 during 2002-2006. These decreases in partial recruitment at age are in line with expectations given recent changes in commercial and recreational fishery regulations. For these reasons, summer flounder are currently considered to be fully recruited to the fisheries at age 3, and fully recruited fishing mortality is expressed as the unweighted average of fishing mortality at age for ages 3 to 5. Fishing mortality calculated from the average of the currently fully recruited ages (3-5) varied between 0.94 and 2.13 during 1982-1997, then declined substantially and was estimated to be 0.44 in 2006 (Figure 30).

Summer flounder spawn in the late autumn and early winter (peak spawning on November 1), and age 0 fish recruit to the fishery during the autumn after they are spawned. For example, summer flounder spawned in autumn 1987 (from the November 1, 1987 spawning stock biomass) recruit to the fishery in autumn 1988, and appear in VPA tables as age 0 fish in 1988. The F08_BASE run indicates that the 1982 and 1983 year classes were the largest of the VPA series, at 73 and 79 million fish, respectively. The 1988 year class was the smallest of the series, at only 13 million fish. The 2006 year class is estimated at 31 million fish, below the time series average of 36 million (Figure 29). Spawning stock biomass (SSB; Age 0+) declined 71% between 1983 and 1989 (23,900 mt to 6,900 mt), but has increased six-fold to 41,700 mt in 2005, before falling slightly to 40,500 mt in 2006 (Figure 28).

Retrospective analysis of the summer flounder ADAPT VPA F08_BASE run was carried out for terminal catch years 1995-2006. The retrospective analysis indicates a pattern of overestimation of fully recruited F (ages 3-5) for 1995-1997, followed by a pattern of underestimation of F for 1998-2005, continuing the pattern observed in the last several assessments (NEFSC 2000, MAFMC 2001a, NEFSC 2002, NEFSC 2005, Terceiro 2006, S&T 2006)(Figure 31). For the last three years, fishing mortality was underestimated by 42% for 2003, by 25% for 2004, and by 15% for 2005, relative to the terminal year 2006 estimates. Spawning stock biomass has been generally been overestimated in the last 3 years, by 54% for 2003, 23% for 2004, and 7% for 2005, relative to the terminal year 2006 estimates (Figure 32). There is no consistent retrospective pattern in the estimation of the abundance of age 0 fish over the last three years (Figure 33).

As the previously accepted, peer-reviewed assessment model, the 2006 NMFS S&T ADAPT VPA (Terceiro 2006) has been updated through 2007. The updated run (VPA_2007) exhibited the same time series trends and retrospective characteristics as the 2006 NMFS S&T run. Using the ADAPT VPA model assuming constant M = 0.20, the stock is overfished and overfishing is occurring when compared to existing BRPs (ADAPT VPA F2007 = 0.311, 11% above the existing F BRP = Fmax = FMSY = 0.280; ADAPT VPA SSB2007 = 42,123 mt, 47% of the existing SSB BRP = SSBMSY = 89,411 mt).

4.3 Age Structured Assessment Program (ASAP)

Fishing mortality rates and stock sizes were estimated using the ASAP model as implemented in the NOAA Fisheries Toolbox (NFT) ASAP version 2.0.9. The catch at age, mean weights at age, maturity at age, and survey index calibration time series were input as in the ADAPT VPA F08_BASE run (see section 4.2.1). An instantaneous natural mortality rate of M = 0.2 was assumed for all ages in all years. Fishery selectivities (partial recruitment) were estimated either at each age or by fitting a single (flat-topped) logistic curve. Winter, spring, and mid-year survey indices and all survey recruitment (age-0) indices were compared to population numbers of the same age at the beginning of the same year. Fall survey indices were compared to population numbers one year older at the beginning of the next year. In developing the ASAP2 F08_BASE run, lognormal error distributions were assumed for the total catch in weight, research survey catch at age calibration indices, internal Beverton-Holt stock-recruitment relationship and parameters, selectivity parameters, fishing mortality (Fmult) parameters, survey catchability parameters, and estimated stock numbers at age. A multinomial distribution was assumed for fishery catch at age. A number of additional initial model settings are required in ASAP, including specification of likelihood component emphasis factors (lambdas), size of deviation factors expressed as standard deviations, and penalty functions for extreme fishing mortality estimates. The settings were left at the default values in the first few runs, and changed as runs were developed.

Initial runs used a) the number of sampled commercial fishery trips (ranging from 30 per year in 1995 to 181 in 2006) as the lambda for the effective fishery age composition sample size (ESS), b) a total catch lambda of 10 and CV of 0.01, c) recruitment deviation lambda of 6.74 and CV of 0.5, d) Fmult lambda of 6.74 and CV of 0.9, e) survey index lambdas of 0.4 and CV of 0.9 for all indices and catchability deviations lambda of 10000 (i.e., constant) and CV of 0.9, f) s-r function lambdas of 6.74 and CV of 0.9, and g) N in year 1 lambda of 6.74 with CV of 0.9.

Run configurations adjusting the ESS, lambdas, survey CVs, and time blocks for the estimation of the fishery selectivity pattern were tested to judge the sensitivity of the analysis to these settings. The sequence of runs is summarized in Table 115. As a result of these tests, a) the ESS was set at 200, b) a total catch lambda was set at 10 with CV of 0.10, c) recruitment deviation lambda was set at 0.001 with CV of 0.5, d) Fmult lambda was set at 1.0 with CV of 0.9, e) survey index lambdas were set at 1.0 with CVs of NEC Winter = 0.15, NEC Spring = 0.25, and NEC Fall = 0.25, and catchability deviations lambda of 10000 (i.e., constant) and CVs of 0.9, f) s-r function lambdas were set at 0.001 with CVs of 0.9, and g) N in year 1 lambda was set at 1.0 with CV of 0.9. These settings provided a good fit to fishery total catch and age comps, reasonable fit to survey indices, and a smooth transition in F pattern through the selectivity break in 1994-1995, and resulted in run ASAP run SELEX_94_95 which was then renamed BASE for subsequent tuning.

The next steps were to a) include the additional state agency survey indices at age accepted for the base ADAPT VPA formulation, with initial lambdas set at 1.0 with CVs of 0.40, b) set s-r function lambdas and CVs to 0, and c) set the ESS to the numbers estimated in the BASE run. Subsequent tuning adjustments (runs T1 and T2) attempted to even out patterns in the estimated ESS and fishing mortality using changes in the at-age selectivity estimation (Table 116). Run T3 was fit using a single logistic pattern with a break at 1994/1995. This change provided a smoother F pattern, while maintaining the expected transition from full selection at age 2 in the first block (1982-1994) and full selection at age 3 in the second (1995-2006). Run T4 specified the survey CVs to more closely match the true time series means of CVs of the NEFSC and MADMF series (other state agency CVs not available), but this change failed to improve the model fit to the surveys, and degraded the fit to the fishery age compositions. Following guidance received from Ian Stewart of NMFS NWFSC (pers. comm.) and Chris Legault of NMFS NEFSC (pers. comm.), the final tuning step, run T5, increased the survey CVs by 1.5 to 2.0 times, to allow better fits to the survey indices while maintaining fit to the fishery age compositions. Changes in the stock size and fishing mortality estimates between the BASE and T1 through T5 run configurations were small (Figures 34-36). This last tuning configuration, F08_BASE_T5, was used as the final ASAP base run.

The annual selection of age-1 fish decreased from about 0.54 during the first time block of selectivity estimation (1982-1994) to about 0.16 during the second block, 1995-2006. The annual selection of age-2 fish decreased from about 0.97 during the first time block of selectivity estimation (1982-1994) to about 0.72 during the second block, 1995-2006. These decreases in selection at age are in line with expectations given changes in commercial and recreational fishery regulations. For these reasons, summer flounder are currently considered to be fully recruited to the fisheries at age 3, and fully recruited fishing mortality is expressed as the unweighted average of fishing mortality at age for ages 3 to 5. Fishing mortality calculated from the average of the currently fully recruited ages (3-5) varied between 1.04 and 1.93 during 1982-1997, then declined substantially and was estimated to be 0.38 in 2006 (Figure 36).

The F08_BASE_T5 run indicates that the 1983 year class was the largest of the series, at 83 million fish. The 1988 year class was the smallest of the series, at only 12 million fish. The 2006 year class is estimated at 34 million fish, below the time series average of 36 million (Figure 35). Spawning stock biomass (SSB; Age 0+) declined 73% between 1983 and 1989 (23,300 mt to 6,300 mt), but increased six-fold to 39,900 mt in 2003, before falling to 38,600 mt in 2006 (Figure 34).

Retrospective analysis of the summer flounder ASAP F08_BASE_T5 run was carried out for terminal catch years 1997-2006 – earlier terminal years were not advisable due to the constraints of the selectivity blocks (2-3 years are recommended in each block; Chris Legault, NMFS NEFSC, pers.comm.). The retrospective analysis indicates a pattern of underestimation of fully recruited F (ages 3-5) for 1997-2005 (Figure 37). For the last three years, fishing mortality was underestimated by 38% for 2003, by 32% for 2004, and by 9% for 2005, relative to the terminal year 2006 estimates. Spawning stock biomass has been generally been overestimated in the last 3 years, by 60% for 2003, 44% for 2004, and 9% for 2005, relative to the terminal year 2006 estimates (Figure 38). There is no consistent retrospective pattern in the estimation of the abundance of age 0 fish since 1997 (Figure 39).

4.4 Stock Synthesis 2 (SS2)

Fishing mortality rates and stock sizes were estimated using the Stock Synthesis 2 (SS2) model as implemented in the NOAA Fisheries Toolbox (NFT) SS2 version 2.00o. The catch at age, maturity at age, and survey index calibration time series were input as in the ADAPT VPA F08_BASE run (see section 4.2.1). For the population biology component of the SS2 model runs, growth patterns were estimated for combined sexes from NEFSC survey biological data for the period from 1992-2007, age structure was set at ages 0-15, and length structure at lengths 10 to 79 cm. Growth rates were estimated as of January 1. An instantaneous natural mortality rate of M = 0.2 was assumed for all ages in all years. Fishery selectivity (partial recruitment) was generally estimated by fitting a single (flat-topped) logistic curve, although dome-shaped patterns were explored. Winter, spring, and mid-year survey indices and all survey recruitment (age-0) indices were compared to population numbers of the same age at the beginning of the same year. Fall survey indices were compared to population numbers one year older at the beginning of the next year. In developing the SS2 F08_BASE run, lognormal error distributions were assumed for the total catch in weight, research survey catch at age calibration indices, internal Beverton-Holt stock-recruitment parameters, selectivity parameters, fishing mortality parameters, survey catchability parameters, and estimated stock numbers at age. When modeled with lognormal error, survey selectivity was specified as S = 1 for the relevant age or series of ages indexed by the survey. Normal error distributions were assumed for mean length at age and mean weight at age. A multinomial distribution was assumed for fishery catch at age.

A number of additional initial model settings are required in SS2, including specification of likelihood component emphasis factors (lambdas), size of deviation factors expressed as standard deviations, and penalty functions for extreme fishing mortality estimates. The lambdas were set at 1 for the likelihood components that were intended to influence model fit. Initial runs used a) the number of sampled commercial fishery trips (ranging from 30 per year in 1995 to 181 in 2006) as the lambda for the effective fishery age composition sample size (ESS), b) total catch lambda equal to the default value of 1, and c) initial CVs of 0.15 for the NEFSC Winter survey indices and 0.25 for the NEFSC Spring and Fall indices.

Initial runs used a single fishery catch at age matrix and NEFSC survey indices, and adjusted the number and timing of fishery selectivity blocks and the number of selectivity parameters estimated to judge the sensitivity of the analysis to these settings. The sequence of initial runs is summarized in Table 117.

These tests found that attempts to fit 3 parameters to describe a logistic selectivity pattern for the fishery generally resulted in gradients (a model fit diagnostic) that were too high. Fitting a single pattern for the entire 1982-2006 time series also generally provided gradients that were too high, or poor fits to the fishery age composition for some years in the early part of the time series. Initial runs with 3 time blocks and 1 or 2 selectivity parameters were judged to fit best. A second round of tests showed that the selectivity parameters for the first 2 periods were nearly the same and so subsequent runs included only 2 blocks: 1982-1994 and 1995-2006. Further testing indicated best fits for runs with 2 logistic selectivity parameters (SELEX pattern 20, parameters 1 (peak) and 3 (ascending limb width)), with a total catch weight lambda of 10 (Table 118). A third round of testing incorporated Working Group (WG) recommendations (Mark Maunder, Quantitative Resource Assessment LLC, pers. comm.) for general SS2 model settings, including estimation of R0, R1 deviations, testing of the start date for recruitment deviations, and testing of the effect of the s-r function lambda. These tests are summarized in Table 119., and provided a base configuration (F08_BASE) that would be subject to “tuning” by adjusting the observed ESS and survey index CVs.

The next steps were to include the additional state agency survey indices at age accepted for the base ADAPT VPA formulation and adjust likelihood component lambdas and survey index CVs. Survey CVs were re-set to 0.16 for the NEFSC Winter survey indices, 0.21 for the NEFSC Spring indices, 0.31 for the NEFSC Fall indices, 0.21 for the MADMF Spring and Fall indices (based on the average of the annual aggregate indices), and 0.40 for the CTDEP, RIDFW, NJDFW, DEDFW, MDDNR, VIMS, and NCDFW indices (no average values available; based on expectation that state survey indices at

ages would be less precise than NEFSC surveys). The s-r function lambdas and CVs were set to 0 to remove the influence on the fit (as in the ASAP model building exercise), but allow internal estimation of the s-r function and reference points. The effects of these changes are summarized in Table 120 (run F08_BASE).
Following guidance received from Ian Stewart of NMFS NWFSC (pers. comm.) and Chris Legault of NMFS NEFSC (pers. comm.), the next step was to use the estimated ESS to observed sample size (OSS) ratio for the times series (1.863) to "tune" the model to the fishery age compositions by multiplying the input observed ESS by this ratio (run F08_BASE_T1). A recommended second tuning step was to increase the survey CVs by 1.5 to 2.0 times (variance adjustment in absolute terms of +0.14, +0.19, and +0.20), to allow better fits to the survey indices while maintaining fit to the fishery age compositions (run F08_BASE_T2). The ESS/OSS ratio tuning step had a greater effect on the results than did increasing the survey CVs (Table 120). Changes in the stock size and fishing mortality estimates between the BASE, T1 and T2 run configurations are presented in Figures 40-42. This last tuning configuration, F08_BASE_T2, was used as the final SS2 base run.

The annual selection of age-1 fish decreased from about 0.51 during the first time block of selectivity estimation (1982-1994) to about 0.15 during the second block, 1995-2006. The annual selection of age-2 fish decreased from about 1.00 during the first time block of selectivity estimation (1982-1994) to about 0.68 during the second block, 1995-2006. These decreases in selection at age are very similar to those estimated for the same time blocks using the ASAP model (see section 4.3), and are in line with expectations given changes in commercial and recreational fishery regulations. For these reasons, summer flounder are currently considered to be fully recruited to the fisheries at age 3, and fully recruited fishing mortality is expressed as the unweighted average of fishing mortality at age for ages 3 to 5. Fishing mortality calculated from the average of fully recruited ages 3-5 varied between 1.14 and 1.84 during 1982-1996, then declined substantially and was estimated to be 0.43 in 2006 (Figure 42).

The F08_BASE_T2 run indicates that the 1983 year class was the largest of the series, at 64 million fish. The 1988 year class was the smallest of the series, at only 10 million fish. The 2006 year class is estimated at 32 million fish, equal to the time series average of 32 million (Figure 41). Spawning stock biomass (SSB; Age 0+) declined 68% between 1983 and 1990 (28,900 mt to 9,200 mt), but increased four-fold to 39,200 mt in 2005, before falling to 38,800 mt in 2006 (Figure 40).

Retrospective analysis of the summer flounder SS2 F08_BASE_T2 run was carried out for terminal catch years 1997-2006 – earlier terminal years were not advisable due to the constraints of the selectivity blocks (2-3 years are recommended in each block; Ian Stewart, NMFS NWFSC, pers. comm.; Chris Legault, NMFS NEFSC, pers. comm.). The retrospective analysis indicates a pattern of underestimation of fully recruited F (ages 3-5) for 1997-2005 (Figure 43). For the last three years, fishing mortality was underestimated by 42% for 2003, by 37% for 2004, and by 12% for 2005, relative to the terminal year 2006 estimates. Spawning stock biomass has been generally been overestimated in the last 3 years, by 51% for 2003, 38% for 2004, and 8% for 2005, relative to the terminal year 2006 estimates (Figure 44). There is no consistent retrospective pattern in the estimation of the abundance of age 0 fish since 1997 (Figure 45).

4.5 Considerations for Model Selection

4.5.1 Comparative BASE Model Results

Fishing mortality rates in 2006 and stock sizes in 2007 were estimated using a) the ADAPT method for calibration of the VPA (Parrack 1986, Gavaris 1988, Conser and Powers 1990) as implemented in the NOAA Fisheries Toolbox (NFT) ADAPT VPA version 2.7.7, b) the ASAP model as implemented in the NOAA Fisheries Toolbox (NFT) ASAP version 2.0.13, and c) the Stock Synthesis 2 (SS2) model as implemented in the NOAA Fisheries Toolbox (NFT) SS2 version 2.00o. The catch at age, maturity at age, and survey index calibration time series were input as in the ADAPT VPA F08_BASE run (see section 4.2.1), with fishery and survey data through 2006 (fishery) and 2007 (selected surveys) for ADAPT and through 2006 for ASAP and SS2. The ADAPT VPA was used as the platform to select the base set of indices to carry forward in the base case for all 3 models because the existing NFT ADAPT software has very useful diagnostic features for judging the calibration performance of the indices. The final base run configurations include 39 survey indices at age.

The ASAP and SS2 models require many additional assumptions and model settings as compared to the ADAPT VPA, including setting for emphasis factors (lambdas) and measures of deviation for the catch, survey and fishery age composition, and s-r function likelihood components. The base case ASAP and SS2 model configurations were developed in parallel fashion, although the experience gained in first developing the SS2 base case helped guide the ASAP base case formulation, especially with regards to the timing of selectivity blocks. Likewise, recommendations provided on the “tuning” of the SS2 model were adopted to guide the “tuning” of the ASAP model. Details are provided in section 4.3 for ASAP and section 4.4 for SS2. Figures 46-48 compare the estimates of Spawning Stock Biomass (SSB; mt), recruitment at age 0 (R; 000s) and fully-recruited fishing mortality rate (F, ages 3-5) from the base case configurations for the ADAPT VPA, ASAP, and SS2 models. In general, the 3 models provided similar results in terms of both the trend and current estimates of fishing mortality and stock size. The SS2 base case provided lower estimates of recruitment in the early years of the time series compared to the ADAPT VPA and ASAP models.

Retrospective analysis of the 3 models was carried out for terminal catch years 1995-2006 for the ADAPT VPA, and for 1997-2006 for ASAP and SS2 (1995-1996 were omitted to avoid estimation problems related to the selectivity block break between 1994 and 1995). For the ADAPT VPA over the last 3 years, fishing mortality was underestimated by 42% for 2003, by 25% for 2004, and by 15% for 2005, relative to the terminal year 2006 estimates. ADAPT VPA estimates of SSB were overestimated in the last 3 years, by 54% for 2003, 23% for 2004, and 7% for 2005, relative to the terminal year 2006 estimates. For ASAP over the last 3 years, fishing mortality was underestimated by 37% for 2003, by 32% for 2004, and by 9% for 2005, relative to the terminal year 2006 estimates. ASAP estimates of SSB were overestimated in the last 3 years, by 60% for 2003, 44% for 2004, and 9% for 2005, relative to the terminal year 2006 estimates. For SS2 over the last 3 years, fishing mortality was underestimated by 41% for 2003, by 37% for 2004, and by 12% for 2005, relative to the terminal year 2006 estimates. SS2 estimates of SSB were overestimated in the last 3 years, by 51% for 2003, 38% for 2004, and 8% for 2005, relative to the terminal year 2006 estimates. There was no consistent retrospective pattern in the estimation of the abundance of age 0 fish over the last three years for any of the 3 models.

Table 121 presents the calculated Mohn's rho diagnostic ([retro year estimate – current year estimate]/current year estimate) for the 3 models for SSB and F. If the cumulative value of rho is used to judge the performance of the model over the retrospective time interval, then the ASAP model performed best (smallest cumulative sum), with the ADAPT VPA and SS2 exhibiting more severe retrospective error in fishing mortality estimates (comparable larger cumulative sums). The ASAP model also exhibited the smallest retrospective error in SSB, followed by the SS2 and ADAPT VPA models. See the individual model sections (4.2-4.4) for the retrospective plots of F and SSB for each model.

Comparative BASE Model Characteristics (Pros and Cons)

ADAPT VPA

Pros: a) relatively simple model compared to the SCAA models, results dictated by the input data, within the constraints of the catch equation formulation, b) well developed bootstrap routine to estimate uncertainty of current year estimates, c) well developed interface with AGEPRO stochastic projection software.

Cons: a) current implementation lacks flexibility to model multiple fleets, multiple sexes, multiple areas, different selectivity assumptions, b) exhibits most severe retrospective error for the summer flounder data.

ASAP

Pros: a) current implementation is a moderately complex, flexible SCAA capable of modeling multiple fleets and multiple approaches to survey index modeling, b) well developed MCMC routine to estimate uncertainty of current year estimates, c) well developed interface with AGEPRO stochastic projection software, d) exhibits least severe retrospective error for the summer flounder data.

Cons: a) current implementation lacks flexibility to model multiple sexes or multiple areas.

SS2

Pros: a) current implementation is a very complex, flexible SCAA capable of modeling multiple fleets, multiple sexes, multiple areas, and multiple approaches to survey index modeling.

Cons: a) exhibits relatively severe retrospective error for the summer flounder data, b) current MCMC implementation lacks an interface with AGEPRO stochastic projection software.

4.5.2 Alternative ASAP and SS2 Model Configurations

ASAP

Alternative configurations of the ASAP v2.0.13 model were tested to investigate the sensitivity of the qualitative assessment conclusions to different ways of modeling the BASE case assessment data. In the BASE case, the fishery catch data are modeled as a single aggregate catch at age matrix, with multinomial error distribution, and a single, time-varying fishery selectivity pattern is estimated for the combined fisheries. In the BASE case, the survey indices at age are modeled as individual indices at age with lognormal error, and survey selectivity specified as S = 1 for the relevant age or series of ages indexed by the survey.

In the first alternative configuration (F08_MULTI), the six fishery catch at age components (NER [ME-VA] commercial landings, NC commercial landings, commercial trawl fishery discards, commercial scallop dredge fishery discards, recreational fishery landings, and recreational fishery discards) were modeled separately, each with a multinomial error assumption. Flat-topped (single logistic, asymptotic) time-varying selectivity patterns were modeled for landings; dome-shaped patterns were modeled for the discards. For the trawl and scallop discards the By-Age selectivity model, with S = 1 fixed at age 1 for the trawl fishery discards and S = 1 fixed at age 2 for the scallop fishery discards, was used to model a dome-shaped pattern that was constant for the time series. For the recreational discards, a time-varying double logistic model was used to model a dome-shaped pattern. In the F08_MULTI configuration, survey indices were modeled as in the BASE case.

In the second alternative configuration (F08_SVAge comp), the survey indices at age were modeled with a multinomial error assumption when feasible (e.g., NEFSC Winter survey indices at age, ages 1-7+). A constant, flat-topped selectivity pattern was used for surveys with a full range of ages. The NEFSC Winter, Spring and Fall, CTDEP Fall, RI Monthly, and NJDFW Monthly survey indices at age were modeled in this manner. The MADMF Spring and Fall, CTDEP Spring, and RIDFW Fall surveys were modeled using the By-Age model for the selected ages included in the BASE case tuning set. The stand-alone age 0 recruitment index series (DEDFW Inland, MDDNR and VIMS) were modeled as in the BASE case (lognormal, S= 1). In the F08_SVAgecomp configuration, the single fishery catch at matrix was modeled as in the BASE case.

In the third configuration (F08_MULTI_SVAGE), the F08_MULTI and F08_SVAgecomp configurations were combined, so that the 6 fishery catch at age components were modeled as in the MULTI configuration, while the survey indices at age were modeled as in the SVAgecomp configuration.

Table 122 summarizes the run diagnostics and results for the three ASAP alternative runs. No "tuning" steps were undertaken for the three ASAP alternatives. The diagnostics for the SVAgecomp configuration was acceptable. The MULTI and MULTI_SVAGE configurations both exhibited problems in the commercial trawl and scallop fishery discard selectivity fits and S-R function parameter estimation (due to very high recruitment estimates at low SSB). The multiple fisheries ASAP configuration results were sensitive to the method of selectivity for the commercial discards, and the estimated age compositions did not match the observed discard at age well. If may not be feasible to model the trawl and scallop fishery discards as separate fleets. Alternatively, future work might consider methods to extend the commercial fishery discard series back to 1982.

None of the three ASAP alternatives provided solutions that significantly reduced that pattern of positive residuals in aggregate indices or indices at age (mainly ages 3-5) during the late 1990s/early 2000s that was apparent in the BASE case ADAPT VPA, ASAP, and SS2 models.

The SVAgecomp configuration estimates most closely matched the ASAP BASE case. The MULTI_SVAGE configuration usually provided the most variable and/or highest F estimates when compared to the BASE case and two other alternative configurations. Figures 49-51 compare the estimates of Spawning Stock Biomass (SSB; mt), recruitment at age 0 (R; 000s) and fully-recruited fishing mortality rate (F, ages 3-5) from the ASAP BASE case configuration (F08_BASE_T5) with the three ASAP alternatives.

SS2

Alternative configurations of the SS2 model were tested to investigate the sensitivity of the qualitative assessment conclusions to different ways of modeling the BASE case assessment data. In the BASE case, the fishery catch data are modeled as a single aggregate catch at age matrix, with multinomial error distribution, and a single, time-varying fishery selectivity pattern is estimated for the combined fisheries. In the BASE case, the survey indices at age are modeled as individual indices at age with lognormal error, and survey selectivity specified as S = 1 for the relevant age or series of ages indexed by the survey.

In the first alternative configuration (F08_MULTI), the six fishery catch at age components were modeled separately, each with a multinomial error assumption and a time-varying selectivity pattern. Flat-topped (single logistic, asymptotic) selectivity patterns were modeled for landings; dome-shaped patterns were modeled for the discards. In the F08_MULTI configuration, survey indices were modeled as in the BASE case.

In the second alternative configuration (F08_SVAge comp), the survey indices at age were modeled with a multinomial error assumption (e.g., NEFSC Winter survey indices at age, ages 1-7+) and a constant, flat-topped selectivity pattern. The NEFSC Winter, Spring and Fall, MADMF Spring and Fall, CTDEP Spring and Fall, RIDFW Fall and Monthly, and NJDFW Monthly survey indices at age were modeled in this manner. Stand-alone age 0 recruitment index series (DEDFW Inland, MDDNR and VIMS) were modeled as in the BASE case (lognormal, S= 1). In the F08_SVAgecomp configuration, the single fishery catch at matrix was modeled as in the BASE case.

In the third configuration (F08_MULTI_SVAGE), the F08_MULTI and F08_SVAgecomp configurations were combined, so that the 6 fishery catch at age components were modeled as in the MULTI configuration, while the survey indices at age were modeled as in the SVAgecomp configuration.

Table 123 summarizes the run diagnostics and results for the three alternative SS2 runs. No "tuning" steps were undertaken for the three SS2 alternatives. The diagnostics for all three alternatives were generally acceptable. However, some parameter bounds (constraints) were hit during fitting of the selectivity patterns, and these parameters would need to be fixed near these bounds if one of the alternatives were accepted as the final assessment model run. For the commercial discard fleets modeled with a dome, the selection for the older ages (beyond the age range of the input catch, ages 8 and older) appeared infeasible for both time blocks (either near S = 1 for trawls, or near S = 0.50 for scallop dredges). If may not be feasible to model the trawl and scallop fishery discards as separate fleets. Alternatively, future work might consider methods to extend the commercial fishery discard series back to 1982.

None of the three SS2 alternatives provided solutions that significantly reduced that pattern of positive residuals in aggregate indices or indices at age (mainly ages 3-5) during the late 1990s/early 2000s that was apparent in the BASE case ADAPT VPA, ASAP, and SS2 models. The fishery selectivity pattern for 1995-2006 estimated in the F08_SVAgecomp configuration was "steeper," with higher selectivity at age 2 (S = 0.85) and full recruitment (S = 1.0) by age 3, than in the F08_MULTI and F08_MULTI_SVAGE configurations, for which the landings selectivity patterns tended to have lower selection at age 2 (ranging from 0.30 to 0.65).

Figures 52-54 compare the estimates of Spawning Stock Biomass (SSB; mt), recruitment at age 0 (R; 000s) and fully-recruited fishing mortality rate (F, ages 3-5) from the SS2 BASE case configuration (F08_BASE_T2) with the three SS2 alternatives. The alternative configurations provided similar trends in fishing mortality and stock sizes, although the F08_BASE_T2 and F08_MULTI cases tended to estimate lower Fs and higher stock sizes over the time series. The exception was that the F08_MULTI_SVAGE configuration estimated the lowest F and highest stock sizes in 2005 and 2006.

4.5.3 More Alternative SS2 Model Configurations

More alternative configurations of the SS2 model were tested to investigate the sensitivity of the qualitative assessment conclusions to different ways of modeling the BASE case assessment data. In this case, an alternative sex-structured assessment model was developed for summer flounder. The goal of the analysis was to more accurately represent the population dynamics (e.g. include sex-structure) and extract more of the information contained in the data. The stock assessment model was developed using Stock Synthesis II (Rick Methot, NMFS). It is a sex- and age-structured model. The model starts at an exploited stock size in 1976. The initial age-structure in 1976 is parameterized with substantial flexibility and independent of prior catch. Age 15 is used as a plus group for the dynamics. The catch and catch-at-age data is separated into six fisheries: main commercial fishery, North Carolina commercial fishery, commercial discards, scallop trawl discards, recreational, and recreational discards. The three NMFS trawl surveys are used as relative indices of abundance. Sex-specific catch-at-age data are included for the surveys. Combined sex catch-at-age data are included for the fisheries. Age 11 is used as a plus group for the catch-at-age data. Growth rates differ between males and females. Natural mortality is assumed constant over time and age, but can differ between males and females and is estimated in the model. The proportion female at age zero is assumed to be 0.4. All selectivity curves are dome-shaped except for the winter trawl survey. Fishery selectivities are length-based to accommodate different selectivities at age between the sexes due to differences in growth. Survey selectivities are age based, but the same for each sex. The fishery selectivities have different parameters starting in 1995, except the North Carolina fishery for which the new parameters start in 1989. The fishery selectivities have temporal variability to accommodate the changes in management (e.g. minimum legal size) and spatial differences in size of the fish. The MSY quantities are calculated using the age-specific fishing mortality averaged over 2005-2007.

Several sensitivity analyses were conducted to investigate the model assumptions.

- Assuming asymptotic selectivities for all survey and fisheries, except the discard fisheries. (asymptotic)
- Setting the steepness of the Beverton-Holt stock-recruitment relationship to 0.75 (h75)
- Fixing M at 0.2 for both females and males (M0.2)
- Fixing M at 0.2 and 0.3 for females and males, respectively. (M0.2M0.3)
- Fixing M at 0.2 for females and estimating M for males. (M0.2Mest)
- Using age-specific selectivity for the fisheries (Sage)
- Starting the model in 1982 (Start82)
- Using iterative reweighting to estimate the effective sample size for the catch-at-age data and the standard deviations for the surveys.

The results of the basecase model are much more optimistic than the ASAP assessment (Table 124). The estimates of natural mortality are much higher than currently used and the rate for males (0.54) is higher than the rate for females (0.29). The spawning biomass was estimated to have declined during the late 1970's and 1980's and then rebuilt to above the initial level by 2008 (Figure 55). The initial level in 1976 was about 30% of the unfished level, indicating that by 1976 the stock had already been substantially depleted

(Figure MM1, Table 125). The current spawning biomass level is estimated to be above the level that corresponds to MSY and the current fishing mortality is estimated to be below the level that corresponds to MSY.

The MLE estimate of female natural mortality was statistically significantly different from 0.2 (Figure 56, Table 124) and results are dependent on the value of natural mortality (Table 124). The Male natural mortality was consistently 0.25 units higher than the female natural mortality (Figure 56, Table 124). This is presumably due to the information in the sex-specific survey catch-at-age data.

There is substantial scatter in the stock-recruitment estimates and there is no clear evidence of a stock-recruitment relationship (Figure 57). The sensitivity analysis with the steepness of the Beverton-Holt stock-recruitment relationship fixed at 0.75 fits the data significantly worse than the assessment without a stock recruitment relationship (Table 124). The estimates of natural mortality were higher than in the basecase.

The asymptotic selectivity and age based fishery selectivity sensitivity analyses fit the data significantly worse than the basecase, but the estimates of natural mortality are still substantially higher than used in the ASAP assessment and the results more optimistic (Table 124).

The sensitivity analysis that starts in 1982 provides similar results to the basecase (Table 124). The sensitivity that iteratively reweights the catch-at-age sample size and the standard deviations for the survey likelihoods estimates higher values for natural mortality and is generally more optimistic (Table 124).

All the sensitivities, except when the steepness of the stock-recruitment relationship is fixed at 0.75, estimate that the spawning biomass in 2008 is above the value that corresponds to MSY (Table 124). All sensitivities estimate that the current fishing mortality rate is less than that corresponding to MSY. The spawning biomass corresponding to MSY as a fraction of the unfished spawning biomass is low due to the lack of a stock-recruitment relationship (Table 124). Under current levels of fishing, the base case assessment estimates that continuing at that rate would produce a spawning biomass of about 40% of the unfished level (Table 125)

All of the model runs have positive definite hessians, but the maximum gradient was often larger than that used for the stopping criteria. In some cases a local minima was obtained and the model had to be rerun. Usually, the local minimum was obtained at an unrealistic parameter value (e.g. the estimate of natural mortality was unrealistically low).

4.6 Selected Modeling Approach

4.6.1 Model Selection Justification

After considerable debate focused mainly on the assumption for the natural mortality rate (M), use of a single or two-sex model, and the characteristics of fishery selection patterns, the SDWG concluded that the final assessment for 2008 would be conducted using the ASAP model with two fleets (combined landings modeled with a flat-topped [single logistic] pattern and combined discards modeled with a dome-shaped [double logistic] pattern), a single sex using a combined sex vector of M at age, and surveys configured as indices at age with a lognormal error assumption.

A two-fleet configuration was chosen because it allowed the landings and discards selectivity patterns to be modeled separately, judged to be an improvement over the single fleet F08_BASE runs (see Section 4.3). Six-fleet configurations were also considered in both ASAP and SS2 (Section 4.5.2). In ASAP multiple fleet configurations, landings were modeled with flat-topped patterns (single logistic, asymptotic) for two time periods (break between 1994/1995) and discards were modeled with dome-shaped patterns (double logistic or by-age) for the same periods. In a six-fleet single-sex configuration in ASAP, the fishery selectivity patterns were not well estimated as some parameter estimates were constrained by bounds during fitting, and so the SDWG concluded that a six-fleet model in ASAP could not be currently be constructed without fixing some of the parameters. Comparable problems were encountered in a similarly constructed six-fleet single-sex SS2 model (Section 4.5.2). A six-fleet, two-sex model constructed in SS2 (Section 4.5.3) showed reduced estimation problems by modeling selection for the fleets as annual length-based double-logistic (domed) patterns, although some parameters were still constrained by bounds in model fitting. The domed-shaped patterns for both landings and discards fit better than flat-topped patterns in the SS2 six-fleet two-sex model. The SDWG concluded, however, that these strongly domed patterns and the “cryptic biomass” that was implied (biomass generated by the model that has not been observed in either the fishery or surveys) could not be accepted given the lack of supporting data or assumptions external to the model.

The two-sex configuration modeled in SS2 provided the ability to estimate differences in fishing mortality rates and stock size trends by sex (Section 4.5.3). However, the SDWG was unable to determine the effect on assessment results of modeling some of the surveys by sex, some as combined sexes, and the landings and discards as combined sexes. The ability to compile survey indices by sex is currently limited to the NEFSC surveys – much more future work will need to be completed by the SDWG to compile state survey indices by sex (which may be feasible since many of them are aged using NEFSC age-length keys).

A potentially more difficult task will be to re-compile the landings and discards at age by sex, because fishery dependent samples by sex are not available. The SDWG will need to perform future research to determine if a feasible approach can be developed to re-compile the fishery catch by sex.

The ADAPT VPA, ASAP and SS2 models provide comparable results when configured similarly (see Section 4.5.2). ASAP was chosen as the final model framework for the 2008 assessment mainly because ASAP a) provides the capability to model multiple fleets and multiple approaches to survey modeling, b) can incorporate data on changes in growth expressed as annual mean weights and maturity at age, c) exhibits the least severe retrospective pattern in a BASE run comparison among ADAPT VPA, ASAP, and SS2 (see section 4.5.1), d) provides a well developed MCMC routine to estimate the uncertainty of current year estimates and facilitate completion of TOR 5.0, and e) provides a well developed interface with the NOAA NFT AGEPRO stochastic projection software, to facilitate completion of TOR 8.0.

The SDWG assumed a natural mortality rate (M) of 0.20 for females and 0.30 for males based mainly on recently observed maximum ages (tmax) in NEFSC survey data of 14 years (76 cm, in NEFSC Winter Survey 2005) for females and 12 years (63 cm, in NEFSC Spring Survey 2007) for males, and the expectation that larger and older fish are likely if fishing mortality rates are maintained near current rates over the next several years. The assumptions were further guided by a) the 3/tmax (5% surviving to tmax at F = 0; Vetter 1988, Quinn and Deriso 1999) and 4.22/tmax (1.5% surviving to tmax at F = 0; Hewitt and Hoenig 2005) rule-of-thumb approaches, and b) the current SDWG working papers on summer flounder growth and maturity prepared by Brust, Powell, and Wong. A combined sex M-schedule at age was developed for use in ASAP by assuming these initial M rates by sex, an initial proportion of females at age 0 of 0.40 derived from the NEFSC Fall survey indices by age and sex, and population abundance decline over time at the sex specific M rates. The final abundance weighted combined sex M-schedule at age ranged from 0.26 at age 0 to 0.24 at age 7+, with a mean of 0.25. The new assumption for M (changed from M=0.2 for both sexes, all ages) resulted in substantial change in the summer flounder assessment by rescaling to increase estimates of stock size, biomass and fishing mortality rates, when compared to previous assessments and current BASE case runs.

4.6.2 Final ASAP Model with Terminal Year 2006

Subsequent to the April 2008 SDWG meeting, the final ASAP two-fleet single-sex model run with terminal year 2006 was subject to a single tuning step, by revising the input fleet Observed Sample Size for the two fleets by the ESS/OSS ratio, as in SS2 tuning. The ratio for the landings fleet was 0.95, and so no change from the input value of 200 was made. The ratio for the discards fleet was 9.4, and so the input value was increased from 10 to 90. Additional tuning by increasing the input survey CVs was not done at this stage, since tuning of the ASAP F08_BASE case run indicated that the impact of that tuning was minimal. The F08_FINAL_T2006 run exhibits a retrospective pattern in recent years similar to those of the BASE case runs in the other models – underestimation of F and overestimation of SSB, with no strong pattern evident for recruitment at age 0 (Figures 58-60).

Summer flounder stock size (SSB, R) and fishing mortality (F) as estimated by the S&T 2006 ADAPT VPA assessment (one fleet, mean M = 0.2), the ASAP F08_BASE case model run (one fleet, mean M = 0.20), and the F08_FINAL_T2006 run (two fleets, mean M = 0.25) are summarized in Figures 61-63. The three runs provide similar long term trend in stock size and F, with the F08_FINAL_T2006 run providing intermediate results in terms of recent SSB in comparison to the other two, but higher recent levels of recruitment and F.

The F08_FINAL_T2006 run was also compared with sensitivity runs for four alternative specifications of the M-schedule: 1) F1_M3 - female M = 0.10, male M = 0.30, mean = 0.18, 2) F2_M2 - female M = 0.20, male M = 0.20, mean = 0.20, 3) F2_M4 - female M = 0.20, male M = 0.40, mean = 0.29, and 4) F2_M5 -female M = 0.20, male M = 0.50, mean = 0.33. The F08_FINAL_T2006 results are compared with those from the four M alternative runs in Figures 64-66. The rescaling of the stock size and fishing mortality rate estimates as mean M increases or decreases is readily apparent, while time series trends are the same. The F08_FINAL_T2006 results are intermediate with respect to the alternative assumption M runs. The final ASAP model with terminal year 2007 (The F08_FINAL_T2007) is provided in the following section (5.0).

5.0 Based on the "best" model or models, estimate fishing mortality rate, recruitment, spawning stock biomass, and total stock biomass for the current year and characterize the uncertainty of those estimates. If possible, also include estimates for earlier years with uncertainty estimates.

5.1 Final ASAP Model with Terminal Year 2007

The F08_FINAL_T2006 run was updated with fishery catches and research survey indices through 2007 to create run F08_T2007_T1. Input survey CVs were maintained at the same tuning step values as in the F08_FINAL_T2006 run (NEFSC Winter = 0.3, NEFSC Spring = 0.4, NEFSC Fall = 0.6, all State Agency = 0.6). The input Observed Sample Sizes (OSS) were maintained at 200 for the Landings but reset to 10 for the Discards, as in the initial T2006 runs. The F08_T2007_T1 results were then used to calculate the Effective Sample Size (ESS) for the two fleets; the resulting ESS/OSS ratio was then used to revise the input OSS to 173 for the Landings (0.86 ESS/OSS ratio) and to 101 for the Discards (10.10 ESS/OSS ratio) to configure the F08_T2007_T2 run.

Results for the F08_FINAL_T2006, F08_T2007_T1, and F08_T2007_T2 runs are compared in Figures 67-69; the long term trends are very similar. The only significant differences occur between the T2006 and T2007 runs in stock sizes (SSB and R) during 2003-2006. The patterns in fishing mortality rates are nearly identical. Differences between the two T2007 tuning runs are also very minor. The ASAP F08_T2007_T2 configuration was adopted as the final assessment model run.

The F08_T2007_T2 run exhibits a retrospective pattern in recent years similar to those of the BASE case runs in the other models and the F08_FINAL_T2006 run – generally an underestimation of F and overestimation of SSB since 1997, with no strong pattern evident for recruitment at age 0 (Figures 70-72). Over the last 3 years, the annual retrospective change in fishing mortality has ranged from +30% to -5%; the annual retrospective change in SSB has ranged from -29% to +6%; the annual retrospective change in recruitment has ranged from +12% to +44%.

The F08_T2007_T2 run was also compared with sensitivity runs for four alternative specifications of the M-schedule: 1) F1_M3: female M = 0.10, male M = 0.30, mean =

0.18, 2) F2_M2: female M = 0.20, male M = 0.20, mean = 0.20, 3) F2_M4: female M = 0.20, male M = 0.40, mean = 0.29, and 4) F2_M5: female M = 0.20, male M = 0.50, mean = 0.33. The F08_T2007_T2 results are compared with those from the four M alternative runs in Figures 73-75. As with the T2006 runs, the rescaling of the stock size and fishing mortality rate estimates as mean M increases or decreases is readily apparent, while time series trends are the same. The F08_T2007_T2 results are intermediate with respect to the alternative assumption M runs.

Summary estimates for the 2008 assessment final model ASAP F08_T2007_T2 run are provided in Table 126, and population number and fishing mortality estimates at age are provided in Tables 127 and 128. The full report of the F08_T2007_T2 run is provided in the TOR 5 Appendix. Fishing mortality calculated from the average of the currently fully recruited ages (3-7+) was very high, varying between 1.143 and 2.042 during 1982-1996. The fishing mortality rate has declined to below 1.00 since 1996 and was estimated to be 0.288 in 2007 (Figure 76). There is an 80% probability that the fishing mortality rate in 2007 was between 0.253 and 0.325 (Figures 77 and 78). Spawning stock biomass (SSB) declined from 24,674 mt in 1982 to 7,017 in 1989, then increased to 43,932 mt by 2004. SSB was estimated to be 43,363 in 2007 (Figure 79). There is an 80% chance that SSB in 2007 was between 39,325 and 48,122 mt (Figures 80 and 81). The arithmetic average recruitment from 1982 to 2007 is 41.6 million fish at age 0. The 1982 and 1983 year classes are the largest in the VPA time series, at 73.5 and 81.6 million fish; the 1988 year class is the smallest at only 12.8 million fish. The 2007 year class is currently estimated to be about 40.0 million fish (Figure 79).

6.0 Examine and evaluate the role of the environment on past and present summer flounder recruitment success.

The SDWG has prepared two working group papers in support of this term of reference. The complete working group papers (Working Paper 11 and 12) are provided in Appendix 6, with a summary of these papers provided and findings given below.

The first document explored the hypothesis that relatively cold water temperature, or some mechanism associated with cold and/or severe weather, is correlated with poor recruitment success for summer flounder. Therefore, the relationships between water temperature anomalies, NAO indices and metrics of summer flounder recruitment success were examined by applying the general approaches of Brodziak and O'Brien (2005) and Megrey et al. (2005) to summer flounder Recruit-Spawner (RS) data and relevant environmental data. Brodziak and O'Brien (2005) examined relationships between environmental indices and summer flounder Recruit-Spawner Anomalies (RSAs) and found the NAO winter index forward lagged by two years was a significant predictor of summer flounder RS ratios, with positive NAO anomalies (wet and mild winters) correlating with positive RSAs (high recruit survival rate).

Spawning stock biomass (SSB), Recruit-Spawner Anomalies (RSAs), and absolute recruitment estimates (VPA0; as suggested by Megrey et al (2005)), were computed using data from a version of the 2007 assessment update Virtual Population Analysis

(VPA). NEFSC research survey surface and bottom water temperature anomalies for the Mid-Atlantic Bight North (MABN; Nantucket Shoals to Hudson Canyon) and South regions (MABS; Hudson Canyon to Cape Hatteras) were obtained from the NEFSC database and seasonal temperatures anomalies were computed for the two regions for winter/spring (season 1; January-June) and fall (July-December; lagged forward) for both surface and bottom water temperatures. North Atlantic Oscillation (NAO) climate index monthly values were obtained from the University of East Anglia database and winter (December-March) and fall (September- November) indices were computed (contemporary and forward lagged one or two years).

The current work first attempted to identify potentially significant relationships by using correlation analyses among the environmental factors and SSB, RSAs, and absolute estimates of recruitment (VPA0). A Generalized Additive Model (GAM; Hastie and Tibshirani 1990) was then used to model relationships for environmental (predictive) factors initially identified by the correlation analysis (significant at the p = 0.10 level). The GAM approach is a nonparametric regression technique that relaxes error distribution assumptions in modeling the relationships between independent predictive variables and dependent response variables; it was suggested by Daskalov (1999), Megrey et al. (2005), and Brodziak and O'Brien (2005) as an effective tool for modeling biological responses to environmental factors. The initial null predictive model in the GAM framework used smoothing splines with 3 degress of freedom for each predictive factor. Following the procedures suggested by Brodziak and O'Brien (2005), a stepwise model-selection process was applied to eliminate predictive factors from the model if they had a p-value <= 0.20, with the step repeated until only predictive factors with p <= 0.20 were included in the model. Finally, the time series of the environmental factors with best fitting GAM models were used in an exercise to investigate their performance as potential VPA recruitment calibration indices.

Prominent features of the summer flounder absolute recruitment series (VPA0) include the strong year class that recruited in 1983 and the two weak year classes that recruited in 1988 and 2005; the recruit-spawner anomaly (RSA) series exhibited generally positive anomalies before 1996, and the uniformly negative anomalies since. The strong negative RSAs in 1988 and 2005 correspond to the weak absolute magnitude of recruitments (VPA0) in those years. The pattern of relatively low (negative) RSA since 1995 is one that would be expected for a fish stock exhibiting a Beverton-Holt (1957) asymptotic stock-recruitment relationship as that stock grows toward SSBMSY (Terceiro 2006b).

Several of the regional, seasonal temperature anomalies exhibit significant statistical correlation over the time series. However, the initial GAMs related only those factors that exhibited the strongest statistical correlations; either absolute recruitment (VPA0) or recruit-spawner anomaly (RSA) to the Mid-Atlantic North and South region winter-spring bottom temperature anomalies (MAN_BT1 and MAS_BT2) and the fall and winter NAO Climate indices (NAO_FAL, NAO_FAL_1, and NAO_WIN). There were two final GAMs (GAM1 and GAM2) that were developed relating either the RSA or the VPA0 to predictive factors. The final GAM relating RSA to the predictive factors (GAM1) included only the NAO_WIN index as the predictive factor on the x-axis (i.e., p

<= 0.20). Comparison of the observed NAO_WIN index and estimated RSA indicates a positive and fairly strong predicative relationship, consistent with the results of the correlation analysis.

The final GAM relating VPA0 to the predictive environmental factors included the Mid-Atlantic South region winter-spring bottom temperature anomaly (MAS_BT2) and the and the fall and winter NAO Climate indices (NAO_WIN and NAO_FAL_1; i.e., p <= 0.20). The NAO_WIN index emerged as a significant predictive factor for VPA0 in the GAM model, even though the correlation of this factor with VPA0 was not initially identified as significant (r = 0.08, p = 0.72). It should also be noted that the NAO_FAL index failed to be retained in the GAM (i.e., p > 0.20) even though NAO_FAL was significantly correlated with VPA0 (r = 0.34, p = 0.09). Under GAM2, the combined predictive fit of the retained predictive factors (x-axis) on the absolute magnitude of summer flounder recruitment (VPA0; y-axis) characterizes the strong 1983 year class and the weak 1988 and 2005 year classes relatively well. However, the relationship between VPA0 and the individual environmental factors is relatively weak as evidenced by the wide confidence intervals of the predicted VPA0.

The time series of predictive factors from the GAM2 model (NAO_WIN, NAO_FAL_1 and MAS_BT2) were included as indices of age 0 recruitment (VPA0) in three derivative configurations of the summer flounder ADAPT VPA F07_ALL run to investigate their performance as potential calibration indices (i.e., as proxy indices of recruitment). Inclusion of these predictive factors resulted in increases in the magnitude of the MSR for the alternative runs (MSR = Mean Squared Residual = total sum of squared residuals divided by degrees of freedom), indicating that the inclusion of these environmental factors as recruitment calibration indices degraded the overall fit of the VPA. Estimates of the strong 1983 and weak 1988 year classes, estimated in the converged (stabilized) part of the VPA, were unchanged by the inclusion of the environmental factors; however the estimates of the weak 2005 year class increased by up to 30% in the alternative runs and estimates of the average 2006 year class increased by up to 13% in the alternative runs.

In summary, the results of this work suggest there are relationships between commonly measured environmental factors such as regional water temperature anomalies and larger scale climate indices and metrics of summer flounder recruitment success. However, these relationships are no stronger than those currently modeled using research survey indices of abundance. Inclusion of these environmental factors in alternative configurations of the current summer flounder assessment VPA does not significantly change the pattern of the recruitment time series or increase the precision of current recruitment estimates. The inclusion of the environmental factors in other summer flounder population dynamics models would not be expected to improve the reliability of forecasts or biological reference points.

A second working group paper was developed and applied the time series approach of wavelet analysis to identify if a relationship exists between summer flounder spawning stock biomass and recruitment estimates and two climatic signals that are considered significant in affecting oceanographic and estuarine processes in the Mid-Atlantic Bight. The North Atlantic Oscillation (NAO) is closely related to the Arctic Oscillation (AO) and primarily affects temperature; it has a well-described 8-year cycle and indications of a 4-yr periodicity that are superimposed on longer-term trends. The Pacific North American (PNA) has a well-described teleconnection with the El Niño-Southern Oscillation (ENSO) and has a dominant effect on precipitation and, thus, freshwater inflow, in the northeast region. These periodicities are known to profoundly effect estuarine oyster populations, including recruitment and mortality.

Monthly values for the NAO and PNA indices were obtained from the National Weather Service Climate Prediction Center. Spawning stock biomass (SSB) and absolute recruitment estimates (VPA0) were computed using data from a version of the 2007 assessment update Virtual Population Analysis (VPA).

Wavelet analysis was used to resolve localized variations in the strength of a signal (i.e., the wave) within a time series. With this approach, the original time series is decomposed into a time-frequency space, which allows the dominant components (i.e., the wavelets) that make up the wave to be identified. Soniat et al (2006) provide references to source the mathematical details of the technique. Earlier analyses by conducted by the Rutgers University group evaluated the use of a number of mother wavelets (e.g., Paul, Morlet). Comparison of the two mother wavelets show that, for applications of the type that follows, the Morlet wavelet provides adequate time resolution and superior frequency resolution over the results obtained from the Paul wavelet. As a consequence, the Morlet wavelet is used here. Four wavelet analyses were reported as representative of a number of different analyses. Each is a cross-wavelet analysis, equivalent to a cross-correlational analysis, comparing either the NAO or PNA to either the VPA0 or the spawner-recruit (VPA0/SSB) index.

No evidence exists for a relationship between the PNA and summer flounder recruitment. On the other hand, a relationship between the NAO and summer flounder recruitment is strongly supported. The 8-year periodicity, the dominant periodicity in the NAO, is identified as significantly correlated with an 8-year periodicity in the recruitment indices in all analyses. The significance level consistently exceeds $\alpha = 0.05$. No substantive phase shift occurs. The two periodicities are in near-synchrony so that high NAO and high recruitment indices occur more or less simultaneously. In most analyses, a 4-year periodicity also occurs, although sometimes at a weaker level of significance. This interaction is consistently associated with a phase shift between 1995 and 2000. Such phase shifts are frequently associated with substantive long-term changes in population dynamics. However, this periodicity was no longer significant after the long-term trend in the spawner-recruit data was eliminated. This suggests that the interaction of the two time series was primarily associated with subsets of the time series record. A detailed examination of the coherence over the time series suggests that the 4-year periodicity was stronger pre-1995 and post-2000 and that the phase shift was coincident with a decline in the significance of this periodicity during the intervening years.

The NAO is consistently associated with temperature shifts in the North Atlantic. The present analysis suggest that some portion of the variability in summer flounder recruitment since 1982 can be explained by this climate forcer and its expression in changes in the temperature regime experienced by the fish.

7.0 Biological Reference Points

7.1 Update or redefine biological reference points (BRPs; proxies for BMSY and FMSY), taking into account conclusions from earlier assessments and findings from TOR 6 (i.e., recruitment and the environment). Estimate uncertainty in BRPs. Comment on the scientific adequacy of existing and redefined BRPs.

Background

The calculation of biological reference points for summer flounder based on yield per recruit analysis using the Thompson and Bell (1934) model was first detailed in the 1990 Stock Assessment Workshop (SAW) 11 assessment (NEFC 1990). The 1990 analysis estimated that F_{max} = 0.23. In the 1997 SAW 25 assessment (NEFSC 1997), an updated yield per recruit analysis reflecting the partial recruitment pattern and mean weights at age for 1995-1996 estimated that F_{max} = 0.24. The analysis in the Terceiro (1999) assessment, reflecting partial recruitment and mean weights at age for 1997-1998, estimated that F_{max} = 0.263.

The Overfishing Definition Review Panel (Applegate *et al.* 1998) recommended that the Mid-Atlantic Fishery Management Council (MAFMC) base MSY proxy reference points on yield per recruit analysis, and this recommendation was adopted in formulating the FMP Amendment 12 Overfishing Definition (MAFMC 1999). These reference points were based on the 1999 assessment (Terceiro 1999) and followed what would later be described as the "non-parametric approach" (i.e., biomass reference points calculated as the product of biomass per recruit and a reference period recruitment level; NEFSC 2002a). The 1999 assessment indicated that $F_{threshold} = F_{target} = F_{max}$ = 0.263, yield per recruit (Y/R) at F_{max} was 0.55219 kg/recruit, and January 1 Total Stock Biomass per recruit (TSB/R) at F_{max} was 2.8127 kg/recruit. The median number of summer flounder recruits estimated from the 1999 Virtual Population Analysis (VPA) for 1982-1998 was 37.844 million age-0 fish. Based on this median recruitment level, maximum sustainable yield (Y_{max} as a proxy for MSY) was estimated to be 20,897 mt (46 million lbs) at a Total Stock Biomass (TSB_{max} as a proxy for B_{MSY}) of 106,444 mt (235 million lbs). The biomass threshold, one-half TSB_{max} as a proxy for one-half B_{MSY}, was therefore estimated to be 53,222 mt (118 million lbs). The Terceiro (1999) reference points were retained in the 2000 SAW 31 assessment (NEFSC 2000) because of the stability of the input data and resulting biological reference point estimates.

The MAFMC Science and Statistical Committee (SSC) conducted a peer review of the summer flounder Overfishing Definition in concert with the 2001 assessment update (MAFMC 2001a, b). The SSC reviewed six analyses to estimate biological reference points for summer flounder conducted by members of the Atlantic States Marine Fisheries Commission (ASMFC) Summer Flounder Biological Reference Point Working

Group. After considerable discussion, the SSC decided that although the new analyses conducted by the ASMFC Working Group had resulted in a wide range of estimates, they did not provide a reliable alternative set of reference points for summer flounder. The SSC therefore recommended that F_{target} remain F_{max} = 0.263 because a better estimate had not been established by any of the new analyses. The SSC also reviewed the biomass target (B_{MSY}) and threshold (one-half B_{MSY}) components of the Overfishing Definition and concluded that the new analyses did not justify an alternative estimate of the B_{MSY} proxy. The SSC endorsed the recommendations of SAW 31 which stated that "the use of F_{max} as a proxy for F_{MSY} should be reconsidered as more information on the dynamics of growth in relation to biomass and the shape of the stock recruitment function become available" (NEFSC 2000). The SSC agreed that additional years of stock and recruitment data should be collected and encouraged further model development, including model evaluation through simulation studies. They also encouraged the evaluation of alternative proxies for biological reference points that might be more appropriate for an early maturing species like summer flounder and the development and evaluation of management strategies for fisheries where B_{MSY} is unknown. The SSC indicated that as the stock size increases, population dynamic processes that could reflect density dependent mechanisms should be more closely monitored and corresponding analyses should be expanded, i.e., rates of size and age, maturity, fecundity, and egg viability should be closely monitored as potential indicators of compensation at higher stock sizes. Finally, the SSC recommended that potential environmental influences on recruitment, including oceanographic changes and predation mortality, should be reevaluated as additional recruitment data become available. As a result of the SSC peer review (MAFMC 2001a) the Terceiro (1999) reference points were retained in the 2001 stock assessment (MAFMC 2001b). In the review of the 2002 stock assessment (NEFSC 2002b), SAW 35 concluded that revision of the reference points was not warranted at that time due to the continuing stability of the input data and resulting reference point estimates. The Terceiro (1999) reference points were subsequently retained in the 2003 (Terceiro 2003) and 2004 (SDWG 2004) assessment updates.

The biological reference points for summer flounder were next peer-reviewed by the 2005 SAW 41, based on the 2005 assessment update using fishery data through 2004 and research survey data through 2004/2005 (NEFSC 2005). The SAW 41 Review Panel noted that the Beverton-Holt (Beverton and Holt, 1957; Mace and Doonan 1988; BH) model fit the observed stock-recruitment data well, and provided reference points comparable to those derived from a non-parametric (yield and biomass per recruit) approach. The SAW 41 Panel noted, however, that the quantity of observed stock-recruitment data was limited (22 years), and the data during the early part of the time series, when the SSB was at the lowest observed levels, indicated a level of recruitment near the estimated R_{max}, and exerted a high degree of leverage on the estimation of the model parameters. This leverage resulted in a high value (0.984) for the subsequently calculated steepness of the BH curve, outside of the $\pm$ one standard error interval of the estimate for Pleuronectid flatfish (0.8 $\pm$ 0.1) indicated by Myers (1999). The BH model results suggested that summer flounder SSB could fall to very low levels (<2,000 mt) and still produce recruitment near that produced at SSB_{MSY}. The SAW 41 Panel concluded a) that this result might not be reasonable for the long term, given the recent stock-recruitment history of the stock (i.e., production of a very poor year class in 1988), b) the BH model estimated parameters might prove to be sensitive to subsequent additional years of S-R data, especially if they accumulated at higher levels of SSB and recruitment in the near term, and c) the BH model fit might also be sensitive to the magnitude of recently estimated spawning stock and recruitment, given the recent retrospective pattern of overestimation of stock size evident in the assessment. The SAW 41 Panel recognized that the limited time series of observed stock-recruitment data impacted both reference point estimation approaches (non-parametric and parametric stock-recruitment model) in terms of the potential spawning stock biomass and recruitment levels that might be realized from the stock if fished at fishing mortality rates in the 0.2-0.3 range over the long term. Given these concerns, the SAW 41 Panel advised that the BH model estimates were not suitable for use as biological reference points for summer flounder, and recommended continued use of reference points developed using the non-parametric model approach. FMP biological reference points from the 2005 assessment were F_{max} = F_{MSY} = 0.276, Y_{max} = MSY = 19,072 mt (42.0 million lbs), TSB_{max} = B_{MSY} = 92,645 mt (204.2 million lbs), and biomass threshold of 0.5*TSB_{max} = 46,323 mt (102.1 million lbs; NEFSC 2005).

The most recent peer review of biological reference points for summer flounder occurred in 2006 and was conducted by the National Marine Fisheries Service (NMFS) Office of Science and Technology (S&T)(Methot 2006). The 2006 S&T Peer Review recommended using SSB, rather than TSB as in previous assessments, as the metric for the biomass reference point proxy. The product of the mean recruitment (37.010 million fish) and Y/R at F_{max} was 21,444 mt = 47.276 million lbs (current FMP Amendment 12 proxy for MSY); the product of the mean recruitment and SSB/R at F_{max} was 89,411 mt = 197.118 million lbs (current FMP Amendment 12 proxy for $B_{MSY;}$ Terceiro 2006). The 2006 S&T Peer Review Panel (Methot 2006) recommended adoption of these biological reference points from the non-parametric approach for summer flounder, advising:

"The low level of recruitment observed in 2005 is essentially the same as the low 1988 recruitment, so it is within the range of recruitment fluctuation used in calculating the expected time to rebuild this stock. The Panel finds that the most representative approach to calculating BRPs and rebuilding rates would be to use the entire set of recruitments from 1982-2005. The average, not median, of these recruitments should be used for calculation of biological reference points because much of the stock's accumulated biomass comes from the larger recruitments. Random draws from this set of recruitments would provide a probability distribution of rebuilding rates that is consistent with the occasional occurrence of small recruitments (1988 and 2005) and large recruitments (1982-1987). There is no documented and obvious reason why recruitments were higher during 1982-1987. If such recruitment levels become more common as the stock rebuilds, then the stock may rebuild to an even higher level than is currently targeted. If such recruitment levels do not occur during the next few years of the rebuilding, then the rebuilding target may be not be achieved by the target time to rebuild. More precise forecasts than this are not feasible."

Estimation Methodology

The two biological reference point estimation approaches previously used in the 2005 SAW 41 (NEFSC 2005) and 2006 S&T Peer Review (Terceiro 2006) assessments were again applied in the 2008 assessment work, so as to be potentially complementary and supportive and because using both should build confidence in the results. The automatic objective application of either approach is often compromised by lack of sufficient observation on stock and recruitment over a range of biomass to provide suitable contrast. Thus, it is often necessary to extrapolate beyond the range of observation and to infer the shape of the stock-recruit relationship from limited and variable observations (NEFSC 2002a). The 2001 MAFMC SSC review of summer flounder reference points also noted this concern (MAFMC 2001a).

The non-parametric approach was to evaluate various statistical moments (mean, variance, percentiles) of the observed series of recruitment data and apply the estimated spawning stock biomass and yield per recruit associated with common F reference points to derive the implied spawning stock biomass and equilibrium total yield (landings plus discards). The biomass and yield per recruit models were fit using the NOAA Fisheries Toolbox (NFT) YPR version 2.7.2 software (NFT 2008a). The full time series of recruitment during 1982-2007 as estimated by the 2008 assessment final model ASAP F08_T2007_T2 run was used in the yield and spawning stock biomass calculations at fishing mortality reference points, as per the 2006 S&T Peer Review Panel recommendation. The non-parametric approach assumes that compensatory mechanisms such as impaired growth, maturity, or recruit survival are negligible over the range of biomass considered (NEFSC 2002a). Once the F_{max} reference point (i.e., the F_{max} proxy for F_{MSY}) was determined, a long-term (100 year) stochastic projection of stock sizes and catches was done to provide better consistency between the estimated medians of the BRP calculations and shorter-term (e.g., 1-5 year) projections (Legault 2008 MS).

The parametric approach used fitted parametric stock-recruitment models along with yield and spawning biomass per recruit information to calculate MSY-based reference points following the procedure of Sissenwine and Shepherd (1987). Stock-recruitment models were fit using the NFT SRFIT version 6.3 software (NFT 2008b). Since a wide range of models (Beverton-Holt [BH] and Ricker [RK] models, incorporating autoregressive error, and Bayesian priors for various parameters) had been tested in the 2005 SAW 41 work, the current parametric model exercise was limited to the simple Beverton-Holt and Ricker models (Beverton and Holt 1957, Mace and Doonan 1988, Ricker 1954).

For the 2008 assessment, the ASAP F08_T2007_T2 model run provides the basis for the 2008 proposed biological reference points and stock status. Average values of mean weights at age in the catch and stock, maturity schedule, and partial recruitment pattern for the period 2005-2007 were used as input for ages 0-7+ for BRP calculations (Table 129). In previous assessments (NEFSC 2005 and earlier) for older aged fish (ages 8-15) with very limited or missing samples, Gomphertz functions based on younger ages were used to estimate mean weights for the older ages in the BRP calculations. However, the practice of extending the age structure to age 15 and use of Gomphertz weights for the older ages results in inconsistency between the BRP biomass estimates based on long-term stochastic projections and shorter-term (e.g., 1-5 year) projections used for Total Allowable Landings (TAL) calculations (NEFSC 2002a, Legault 2008 MS). Therefore, to increase consistency between these two types of projections, the age range of the BRP and projection calculations as been set at 0-7+, with 8 additional ages (to age 15) included in the plus group calculation of yield and spawning biomass per recruit (NFT 2008a). The mean weight at age for the plus group (age 7+; ages 7 and older) was updated for this assessment in a new way, by using a weighted average of mean weights for ages 7-15 (observed catch weights for ages 7-10; calculated Gomphertz weights for ages 11-15 as estimated from observed ages 0-10) based on the relative proportions at age given a 2007 total mortality rate of 0.55 (mean M = 0.25 + 2007 F = 0.30; this value is coincidently consistent with the proposed F35% proxy for FMSY).

2008 Assessment Biological Reference Points

Summer flounder stock size (SSB, R) and fishing mortality (F) as estimated by the S&T 2006 assessment (ADAPT VPA, terminal year 2005, one fleet, mean M = 0.2), the S&T 2006 ADAPT VPA assessment model updated with current catch data through 2007 (VPA_T2007; terminal year 2007, one fleet, mean M = 0.2), the F08_T2007_T2_M20 run (terminal year 2007, two fleets, mean M = 0.20), and the F08_T2007_T2 run (terminal year 2007, two fleets, mean M = 0.25) are summarized in Figures 82-84. The four runs provide similar long term trends in stock size and F, with the ASAP F08_T2007_T2 run using mean M = 0.25 generally providing higher stock sizes (SSB and R) since 1995 than the S&T 2006 ADAPT VPA, the ADAPT VPA_T2007, and ASAP F08_T2007_T2_M20 runs using mean M = 0.20.

The combined effects of the new assumption for M and the modeling of landings and discards as distinct fleets (which results in a slightly domed-shaped combined fishery selectivity pattern) result in higher estimates of F reference points, lower estimates of MSY, lower estimates of SSB reference points, and improved stock status with respect to both the F and SSB reference points as compared to the S&T 2006 assessment (Tables 129-131). For the 2008 assessment, the ASAP F08_T2007_T2 model run provides the basis for the 2008 proposed biological reference points and evaluation of stock status that follows.

The SDWG concluded that reference points estimated from the parametric approach were suspect, as the Beverton-Holt function steepness parameters were always very near 1.0 (Table 130). The WG considered Fmax, F40%, and F35% (and their corresponding biomass reference points) from the non-parametric approach as candidate proxies for FMSY and BMSY. Fmax has been used in previous assessments as the proxy for FMSY. The current estimate of Fmax using mean M = 0.25 and updated fishery selectivity and mean weights at age is relatively high (0.558) and the YPR to F relationship does not indicate a well defined peak (Figure 85). As a result, there is little gain in YPR (<5%) at fishing mortality rates higher than F35% = 0.310. However, the corresponding decline in SSBR between F35% = 0.310 (~1.48 kg/r) and Fmax = 0.558 (~0.93 kg/r) is about 37%. The WG concluded that F40% = 0.254 and F35% = 0.310 were candidate proxies that provided sufficient YPR (F40% YPR = 92% of Fmax YPR; F35% YPR = 97% of Fmax YPR) to allow for productive fisheries while also providing for substantial SSBR (F40% SSBR = 176% of Fmax SSBR; F35% SSBR = 155% of Fmax SSBR) to buffer against short-term declines in recruitment. The WG proposes that F35% = 0.310 and the associated MSY (13,122 mt) and SSBMSY (60,074 mt) estimates from long-term stochastic projections be adopted as proxies for FMSY and SSBMSY (Table 131). The WG proposes that F40% = 0.254 be adopted as a fishing mortality rate target for management.

7.2 Evaluate current stock status with respect to the existing BRPs, as well as with respect to updated or redefined BRPs (from TOR 7a).

The age-structured assessment model for summer flounder has changed from an ADAPT VPA model to a forward projecting ASAP model. A new value for natural mortality has been adopted, changing from a constant value of M = 0.20 to age- and sex-specific values that result in a mean value of M = 0.25. Biological reference points have therefore also been revised; the proxy for FMSY changed from Fmax to F35%, and F40% is recommended as an Ftarget.

The summer flounder stock is not overfished and overfishing is not occurring relative to the proposed 2008 assessment biological reference points (Table 131, Figure 86). Fishing mortality calculated from the average of the currently fully recruited ages (3-7+) ranged between 1.143 and 2.042 during 1982-1996. The fishing mortality rate has declined to below 1.000 since 1996 and was estimated to be 0.288 in 2007, below the proposed fishing mortality reference point = F35% = FMSY = 0.310. There is an 80% probability that the fishing mortality rate in 2007 was between 0.253 and 0.325. Spawning stock biomass (SSB) declined from 24,674 mt in 1982 to 7,017 in 1989, then increased to 43,932 mt by 2004. SSB was estimated to be 43,363 in 2007, about 72% of the proposed SSB35% = SSBMSY reference point = 60,074 mt. There is an 80% chance that SSB in 2007 was between 39,325 and 48,122 mt.

The previously accepted, peer-reviewed 2006 NMFS S&T ADAPT VPA assessment model (Terceiro 2006) has also been updated through 2007. Using the ADAPT VPA model assuming constant M = 0.20, the stock is overfished and overfishing is occurring when compared to existing BRPs (ADAPT VPA F2007 = 0.311, 11% above the existing F BRP = Fmax = FMSY = 0.280; ADAPT VPA SSB2007 = 42,123 mt, 47% of the existing SSB BRP = SSBMSY = 89,411 mt).

8.0 Stock Projections

8.1, 8.2, and 8.3 Recommend what modeling approaches and data should be used for conducting single and multi-year stock projections, computing TACs or TALs, and measures of uncertainty. If possible, provide numerical examples of short term projections (2-3 years) of biomass and fishing mortality rate, and characterize their uncertainty, under various TAC/F strategies. If possible, compare projected stock status to existing rebuilding or recovery schedules, as appropriate.

Stochastic projections were made to provide forecasts of stock size and catches in 2008-2012 consistent with the proposed biological reference points. The projections do not explicitly account for the recent retrospective pattern in the assessment, as per the 2006 S&T Peer Review advice (Terceiro 2006). The projections assume that recent (2005-2007) patterns of discarding will continue over the time span of the projections. Different patterns that could develop in the future due to different trip and bag limits and fishery closures have not been evaluated. To increase consistency between the proposed reference points and stock projections, the input fishery selectivity pattern, M-pattern, and mean weights at age were configured in the same way (Table 132). One hundred projections were made for each of the 1000 MCMC realizations of 2008 stock sizes from the final assessment ASAP model F08_T2007_T2 run using NFT AGEPRO version 3.1.3 (NFT 2008). Future recruitment at age 0 was generated randomly from a cumulative density function of the F08_T2007_T2 run recruitment series for 1982-2007 (mean recruitment = 41.6 million fish).

If landings in 2008 are 7,153 mt (15.8 million lbs) and discards are 885 mt (2.0 million lbs), the forecast estimates a median (50% probability) F in 2008 = 0.238 and a median SSB on November 1, 2008 of 46,992 mt, above the proposed biomass threshold of one-half SSBMSY = 30,037 mt (Table 133, Figure 86). Fishing at Frebuild = 0.274 in 2009 results in forecast median (50%ile) landings of 9,211 mt (20.3 million lbs); the corresponding 25%ile of landings is 8,653 mt (19.1 million lbs) (Table 133). Continued fishing at Frebuild = 0.274 during 2010-2012 is forecast to rebuild the stock to SSBMSY = 60,074 in 2012 (Figure 87). Fishing at F35% = 0.310 during 2009-2012 is forecast to result in SSB = 56,471 mt in 2012, below the proposed SSBMSY (Figure 88). Fishing at F40% = 0.255 during 2009-2012 is forecast to result in SSB = 62,181 mt in 2012, above the proposed SSBMSY (Figure 89).

9.0 Review, evaluate and report on the status of the Research Recommendations offered in recent SARC reviewed assessments and in the 2006 "Methot" Review.

Major data and analytical needs for future assessments have been identified in the SAW 35 review of the 2002 assessment (NEFSC 2002a), the SDWG assessment updates for 2003 and 2004 (Terceiro 2003; SDWG 2004), the SAW 41 assessment update (NEFSC 2005), the 2006 assessment and S&T peer review (Terceiro 2006a, 2006b; Methot et al. 2006), the SDWG 2007 assessment update, and by the SDWG for this current benchmark assessment (SDWG 2008). Research recommendations "never die", and are retained in these documents until they are addressed (completed). Therefore, these remaining recommendations have been subset as those that have been completed (between the last benchmark and the current assessment), in progress at present or to be addressed (previously identified), and new (identified by the SDWG for this benchmark assessment (SDWG 2008)).

9.1 Completed

9.1.1 2008 SDWG Responses to Summary Findings of the 2006 NMFS Office of Science and Technology Peer Review

1. Retain the non-parametric approach to biological reference points; there is insufficient contrast to estimate Spawner-Recruitment steepness.
The non-parametric reference points have been retained in this 2008 benchmark assessment.

2. For the non-parametric approach, use SSB to track status of the stock. This is a much more accurate proxy for the reproductive potential of the stock and is consistent with current consideration of spawner-recruitment models as possible replacement for the non-parametric approach. The past use of Jan 1 total stock biomass as the measure of reproductive potential over-represents the contribution of age 0 fish.

SSB has been used as the basis to track the status of the stock with respect to the biomass reference point.

3. Use long-term (1982-2005) average body weight-at-age for calculation of biological reference points. The recent downturn in mean weight-at-age is influenced by shifting sex ratio and should only be used for short-term TAL and SSB calculations.

Due to recent trends in the biological characteristics of the stock, the SDWG concluded that short-term (2005-2007) average body weight-at-age was used for calculation of biological reference points and projections in this 2008 benchmark assessment.

4. Discount the recent downtrend in recruits per spawner. Such a trend is exactly what is expected from near constant recruitment and reduced fishing mortality which allows more spawning biomass per recruit. Further declines are expected as the stock approaches the rebuilt level.

5. Use the arithmetic mean (not median) of long-term (1982-2005) recruitment as the basis for the average level of recruitment expected from a rebuilt stock. Although the five highest recruitments in this time period occur in the first five years, there is no reason to discount the occasional occurrence of such recruitment levels from a rebuilt stock. Median recruitment underestimates the level of biomass expected from a rebuilt stock because most biomass comes from the larger recruitments.

With respect to Findings 4 & 5: as recommended, the recent downturn in R/SSB was discounted in the calculation of reference points in this 2008 benchmark assessment. The arithmetic mean of long-term (1982-2007) recruitment was used as the basis for the long term level of recruitment expected from a rebuilt stock in calculation of the reference points.

6. Revise the survey input to the VPA model so that observations of zero are replaced with a small positive value. This VPA model, as with most assessment models, fits to the logarithm of the observations so cannot explicitly deal with observations of zero. However, the current VPA practice of treating these observations as missing values is probably underestimating the degree to which the stock has rebuilt since the low level in 1990.

As recommended, survey inputs to the VPA model with zero values were replaced with small positive values in the 2007 update. Since then, work performed by the SDWG and Groundfish Assessment Review Meeting (GARM) working groups has indicated that zeros should not be filled with a small value, and so this practice (retain zeros – treat as missing values) has been re-instituted in this 2008 benchmark assessment. The assessment model is now ASAP.

7. Do not make an explicit adjustment for the retrospective pattern in the VPA results. The pattern diminishes in the last year, its cause is not clear, and past patterns in the opposite direction have also diminished after a few years. The several survey indices included in the model increased greatly during the late 1990s and the indices of the oldest age groups have continued to increase. The current model does not track these changes closely, so exploration of alternative models and data interpretations that better reconcile this recent pattern should be a higher priority than the retrospective pattern.

As recommended, no explicit adjustment for the retrospective pattern was made to the 2008 assessment model results or projections.

9.1.2 Other Completed Research Recommendations

1) Evaluate use of a forward calculating age-structured model for comparison with VPA. Forward models would facilitate use of expanding age/sex structure and allow inclusion of historical data. If sex-specific assessments are explored, the implications on YPR should also be investigated.

SDWG Response: This recommendation was addressed for the current (June 2008) benchmark assessment through modeling exercises utilizing the forward projecting models ASAP and SS2.

2) Evaluate trends in the regional components of the NEFSC surveys and contrast with the state surveys that potentially index components of the stock.

SDWG Response: This recommendation was addressed for the current (June 2008) benchmark assessment through TOR 2.1 which examined the potential for an integrated index approach and TOR 3.2 which examined regional difference in the CAA data.

3) Explore statistical methods to develop “combined” survey abundance indices (by age if possible) from state agency survey data, for use in calibration of analytical assessment models.

SDWG Response: This recommendation was addressed for the current (June 2008) benchmark assessment through TOR 2.1 which examined the potential for an integrated index approach.

4) Consider examining alternative expansion factors (i.e, summer flounder landings, all species landings) for discards and subsequent effect on retrospective pattern.
SDWG Response: This recommendation was addressed for the current (June 2008) benchmark assessment under TOR 1.0, although the conclusions of that working paper (Estimation of Commercial Fishery Discards of Summer Flounder: Update 2007 or Revise the 1989-2007 Time Series) recommend no change to the current methodology and additional research into the merits of other estimation methods.

5) Consider treating discards as a separate catch-at-age component, once the summer flounder assessment is implemented in a statistical catch at age framework.

SDWG Response: This recommendation was addressed for the current (June 2008) benchmark assessment through modeling exercises utilizing the forward projecting models ASAP and SS2.

6) Present the VTR Party/Charter boat to MRFSS/FHS comparison in more detail, including stratification by state/federal waters for federal permit holders in the Party/Charter sector.

SDWG Response: A comparison of the VTR Party/Charter boat to MRFSS/FHS was presented in greater detail for the current (June 2008), while the latter part of this recommendation still needs to be addressed.

7) Consider alternative weighting schemes to explore the sensitivity of the VPA calibration to perceived survey index outliers.

SDWG Response: This recommendation was addressed for the current (June 2008) benchmark assessment through TOR 2.1 which examined the potential for an integrated index approach.
8) For the maximum age plots, consider comparing the 90th percentile of max age, which may more effectively show time series trends (particularly for the males).

SDWG Response: This recommendation was addressed for the current (June 2008) benchmark assessment through TOR 3.3 and 3.4 which examined summer flounder maximum ages and life history parameters.

3) Explore the sensitivity of the VPA results to separating the summer flounder stock into multiple components.

SDWG Response: This recommendation was addressed for the current (June 2008) benchmark assessment through under TOR 3.1.

9.2 To Be Addressed or In Progress

High

1) Develop a program to annually sample the length and age frequency of summer flounder discards from the recreational fishery.
SDWG Response: To date, ongoing programs are in place only in the MRFSS, MRFSS For-hire survey, ALS, Connecticut (CTDEP Volunteer Anglers), Maryland (MD-DNR Volunteer Anglers), to sample lengths of recreational discards. Progress has been made but more synoptic data are needed including the age-frequency.

2) The SDWG noted that more comprehensive collection of otoliths, for all components of the catch-at-age matrix, needs to be collected on a continuing basis for fish larger than 60 cm (~7 years). The collection of otoliths and the proportion at sex for all of the catch components could provide a better indicator of stock productivity.

SDWG Response: This recommendation has not been addressed and remains an ongoing data collection need.

3) The SDWG recommends that a reference collection of summer flounder scales and otoliths be developed to facilitate future quality control of summer flounder production aging. In addition, a comparison study between scales and otoliths as aging structures for summer flounder should be completed.

SDWG Response: An exchange of aging structures between NEFSC and NCDMF was completed and a report was reviewed by the 2007 SDWG, in response to a 2005 SAW 41 high priority Research Recommendation. The SDWG noted that while the Fall 2006 aging exchange between NC-DMF and the NEFSC indicated that the current level of aging consistency between NC and NEFSC is acceptable, there is a need to conduct and fund these exchanges more frequently, on a schedule consistent with benchmark assessments.

4) The SDWG noted that the observed change in the sex ratio in NEFSC survey samples may result in the SSB estimates not translating as directly to egg production since there are more males proportionally in those older age-categories. Collecting information on overall fecundity for the stock, both egg condition and production may be a better indicator of stock productivity.

SDWG Response: This recommendation has not been addressed and remains an ongoing data collection need.

5) Investigate trends in sex ratios and mean lengths and weights of summer flounder in state agency and federal surveys catches.

SDWG Response: While these trends have not been examined in the state survey catches, these trends were examined in the NEFSC spring, autumn, and winter survey data. Additional work to examine and explain these trends in greater detail should be conducted.

Medium

6) Use NEFSC fishery observer age-length keys for 1994 and later years (as they become available) to supplement NEFSC survey data in aging the commercial fishery discard.

SDWG Response: This recommendation has not been addressed by the SDWG, as the age data are not yet available.

7) Consider use of management strategy evaluation techniques to address the implications of harvest policies that incorporate consideration of retrospective pattern (see ICES Journal of Marine Science issue of May 2007).

SDWG Response: This recommendation has not been addressed by the SDWG.

Low

8) Consider treating scallop closed areas as separate strata in calculations of summer flounder discards in the commercial fisheries.

SDWG Response: This recommendation has not been addressed.

9.3 New

The following major data and analytic needs for future assessments were identified by the SDWG in completing the 2008 June benchmark assessment.

1) Examine the sensitivity of the summer flounder assessment to the various unit stock hypotheses and evaluate spatial aspects of the stock to facilitate sex and spatially-explicit modeling of summer flounder.
2) Conduct further research to examine the predator-prey interactions of summer flounder and other species, including food habitat studies, to better understand the influence of these other factors on the summer flounder population.

3) Evaluate and collect information on the reporting accuracy of recreational discards estimates in the recreational fishery.

4) Examine male female ratio at age-0 and potential factors (i.e. environmental) influence determination of that ratio.

5) Evaluate potential changes in fishery selectivity relative to the fecundity of the stock; analysis should consider the potential influence of the recreational and commercial fisheries.

6) Estimate the sex ratio for all of the catch components.

7) Determine the appropriate level for the steepness of the S-R relationship and investigate how that influences the biological reference points

9.4 Major sources of assessment uncertainty

The SDWG identified the following as ongoing sources of uncertainty in the current assessment.

1) The landings from the commercial fisheries used in this assessment assume no under reporting of summer flounder landings. Therefore, reported landings and associated effort from the commercial fisheries should be considered minimal estimates.

2) The recreational fishery landings and discards used in the assessment are estimates developed from the Marine Recreational Fishery Statistics Survey (MRFSS). While the estimates of summer flounder catch are considered to be among the most reliable produced by the MRFSS, they are subject to error.

3) The length and age composition of the recreational discards are based on data from a limited geographic area (MRFSS, MRFSS For-hire survey, ALS, Connecticut (CTDEP Volunteer Anglers), Maryland (MD-DNR Volunteer Anglers). Sampling of recreational fishery discards on an annual, synoptic basis is needed.

4) The current estimate of M remains an ongoing source of uncertainty.

5) Estimation of the mean weight at age for older fish (i.e. age 10+) remains an ongoing source of uncertainty.
6) The influence of sex specific differences in life history parameters on the assessment model.

ACKNOWLEDGMENTS

Special thanks to Jay Burnett and the staff of the NOAA Fisheries NEFSC Population Biology Branch for their timely preparation of the summer flounder ages used in this assessment.

LITERATURE CITED

[Note: Literature cited in specific working group papers can be found in their respective appendices]

Almeida, F.P., R.E. Castaneda, R. Jesien, R.C. Greenfield, and J.M.Burnett, 1992. Proceedings of the NEFC/ASMFC Summer Flounder, Paralichthys dentatus, Ageing Workshop. NOAA Tech. Memo. NMFS-F/NEC-89. 7p.

Anthony, V. 1982. The calculation of F0.1: a plea for standardization. Northwest Atlantic Fisheries Organization, Serial Document SCR 82/VI/64, Halifax, Canada.

Applegate, A., S. Cadrin, J. Hoenig, C. Moore, S. Murawski, and E. Pikitch. 1998. Evaluation of existing overfishing definitions and recommendations for new overfishing definitions to comply with the Sustainable Fisheries Act. Overfishing Definition Review Panel Final Report. 179 p.

Beverton, R.J.H., and S.J. Holt. 1957. On the dynamics of exploited fish populations. Chapman and Hall, London, facsimile reprint 1993.

Bolz, G., R. Monaghan, K. Lang, R. Gregory, and J. Burnett. 2000. Proceedings of the summer flounder aging workshop, 1-2 February 1999, Woods Hole, MA. NOAA Tech. Memo. NMFS-NE-156. 15 p.

Brodziak, J., and P. Rago. MS 1994. A general approach for short-term stochastic projections in age-structured fisheries assessment methods. Population Dynamics Branch, Northeast Fisheries Science Center, Woods Hole, MA 02543.

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

Bugley, K., and G. Shepherd. 1991. Effect of catch-and-release angling on the survival of black sea bass. N. Amer. J. Fish. Mgmt. 11:468-471.

Burns, T.S., R. Schultz, and B.E. Brown. 1983. The commercial catch sampling program in the northeastern United States. In Doubleday, W.G., and D. Rivard [ed.]. 1983. Sampling commercial catches of marine fish and invertebrates. Can. Spec. Publ. Fish. Aquat. Sci. 66: 290 p.

Clark, S.H. 1979. Application of bottom-trawl survey data to fish stock assessments. Fisheries 4: 9-15

Conser, R.J. and J.E. Powers. 1990. Extension of the ADAPT VPA tuning method designed to facilitate assessment work on tuna and swordfish stocks. Int. Comm. Conserv. Atlantic Tunas, Coll. Vol. Sci. Pap. 32: 461-47.

DeLong, A., K. Sosebee, and S. Cadrin. 1997. Evaluation of vessel logbook data for discard and CPUE estimates. SAW 24 SARC Working Paper Gen 5. 33 p.

Dery, L. M. 1997. Summer flounder, Paralichthys dentatus. In: Almeida, F. P., and T. F. Sheehan, eds. Age determination methods for northwest Atlantic species. http://www.wh.whoi.edu/fbi/age-man.html (February 1997).

Diodati, P.J., and R.A. Richards. 1996. Mortality of striped bass hooked and released in saltwater. Trans. Am. Fish. Soc. 125(2):300-307.

Efron, B. 1982. The jackknife, the bootstrap, and other resampling plans. Phila. Soc. for Ind.and Appl. Math. 38.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. Canadian Atl. Fish. Sci. Adv. Comm. (CAFSAC) Res. Doc. 88/29. 12 p.

Hoenig, J.M. 1983. Empirical use of longevity data to estimate mortality rates. Fish. Bull. 81:898-902.

IPHC. 1988. Annual Report, 1987. International Pacific Halibut Commission. Seattle, Washington. 51 p.

Jones, W.J., and J. M. Quattro. 1999. Genetic structure of summer flounder (Paralichthys dentatus) populations north and south of Cape Hatteras. Marine Biology 133: 129-135.

Kraus, R.T., and J. A. Musick. 2003. A brief interpretation of summer flounder, Paralichthys dentatus, movements and stock structure with new tagging data on juveniles. Mar. Fish. Rev. 63(3): 1-6.

Legault, C. 2008 MS. Setting SSBmsy via stochastic simulation ensures consistency with rebuilding projections. A working paper in support of GARM Reference Points Meeting Term of Reference 4. 8 p.

Lucy, J.A., and T.D. Holton. 1998. Release mortality in Virginia's recreational fishery for summer flounder, Paralichthys dentatus. Virginia Marine Resource Report Number 97-8, 48 p.

Lux, F.E., and L.R. Porter. 1966. Length-weight relation of the summer flounder (Paralichthys dentatus (Linneaus). U.S. Bureau Comm. Fish., Spec. Sci. Rept. Fish., No. 531, 5 p.

Mace, P.M., and I.J. Doonan. 1988. A generalized bioeconomic simulation model for fish population dynamics. N.Z. Fish. Assess. Res. Doc. 88/4.

Malchoff, M.H., and J. Lucy. 1998. Short-term hooking mortality of summer flounder in New York and Virginia. Interim report for Cornell Univ/DEC Project MOU 000024, 6 p.

Megrey, B.A., Y. Lee and S.A. Macklin. 2005. Comparative analysis of statistical tools to identify recruitment-environment relationships and forecast recruitment strength. ICES J. Mar. Sci. 62:1256-1269.

Merson, R.R., C.S. Casey, C. Martinez, B. Soffientino, M. Chandlee, and J.L. Specker. 2000. Oocyte development in summer flounder (Paralichthys dentatus): seasonal changes and steriod correlates. J. Fish. Biol. 57(1): 182-196.

Merson, R.R., M. Terceiro, and J.L. Specker. In review. Length and age at maturity of the summer flounder Paralichthys dentatus (L.). J. Fish. Biol.

Methot, R. 2006. Review of the 2006 Summer Flounder Assessment Update. Chair's Report. NMFS Office of Science and Technology. 6 p.

Mid-Atlantic Fishery Management Council. (MAFMC). 1999. Amendment 12 to the summer flounder, scup, and black sea bass fishery management plan. Dover, DE. 398 p + appendix.

Mid-Atlantic Fishery Management Council. (MAFMC). 2001a. SAW Southern Demersal Working Group 2001 Advisory Report: Summer Flounder. 12 p

Mid-Atlantic Fishery Management Council. (MAFMC). 2001b. SSC Meeting - Overfishing Definition. July 31-August 1, 2001. Baltimore, MD. 10 p.

Myers, R.A., K.G. Bowen, and N.J. Barrowman. 1999. Maximum reproductive rate of fish at low population sizes. Can. J. Fish. Aquat. Sci. 56:2404-2419.

National Research Council (NRC). 2000. Improving the collection, management, and use of marine fisheries data. National Academy Press, Washington, DC. 222 p.

NOAA Fisheries Toolbox (NFT) 2008a. Age Structured Assessment Program (ASAP), version 2.0.17. (Internet address: http://nft.nefsc.noaa.gov).

NOAA Fisheries Toolbox Version 3.0. (NFT). 2008. Age structured projection model (AGEPRO), version 3.1.3 (Internet address: http://nft.nefsc.noaa.gov).

NOAA Fisheries Toolbox Version 3.0. (NFT). 2008a. Yield per recruit program, version 2.7.2. (Internet address: http://nft.nefsc.noaa.gov).

NOAA Fisheries Toolbox Version 3.0 (NFT). 2008. Virtual population analysis program (ADAPT-VPA), version 2.7 (Internet address: http://nft.nefsc.noaa.gov).

NOAA Fisheries Toolbox Version 3.0. (NFT). 2008b. Stock recruitment fitting model, version 6.3 (Internet address: http://nft.nefsc.noaa.gov).

NOAA Fisheries Toolbox Version 3.0. (NFT). 2008. Stock Synthesis 2, version SS2o. (Internet address: http://nft.nefsc.noaa.gov).

Northeast Fisheries Center (NEFC). 1990. Report of the Eleventh NEFC Stock Assessment Workshop Fall 1990. NEFC Ref. Doc. No. 90-09. 121 p.

Northeast Fisheries Science Center (NEFSC). 1993. Report of the 16th Northeast Regional Stock Assessment Workshop (16th SAW). NEFSC Ref. Doc. No. 93-18. 116 p.

Northeast Fisheries Science Center (NEFSC). 1996a. Report of the 20th Northeast Regional Stock Assessment Workshop (20th SAW): Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NEFSC Ref. Doc. No. 95-18. 211 p.

Northeast Fisheries Science Center (NEFSC). 1996b. Report of the 22nd Northeast Regional Stock Assessment Workshop (22nd SAW): Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NEFSC Ref. Doc. No. 96-13. 242 p.

Northeast Fisheries Science Center (NEFSC). 1997a. Report of the 24th Northeast Regional Stock Assessment Workshop (24th SAW): Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NEFSC Ref. Doc. No. 97-12. 291 p.

Northeast Fisheries Science Center (NEFSC). 1997b. Report of the 25th Northeast Regional Stock Assessment Workshop (25th SAW): Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NEFSC Ref. Doc. No. 97-14. 143 p.

Northeast Fisheries Science Center (NEFSC). 2000. Report of the 31st Northeast Regional Stock Assessment Workshop (31st SAW): Stock Assessment Review Committee (SARC) Consensus Summary of Assessments. NEFSC Ref. Doc. No. 00-15. 400 p.

Northeast Fisheries Science Center (NEFSC) 2002. Report of the 35th Northeast Regional Stock Assessment Workshop (35th SAW): SARC Consensus Summary of Assessments. NEFSC Reference Document 02-14. 259 p.

Northeast Fisheries Science Center (NEFSC) 2002a. Final Report of the Working Group on Re-evaluation of Biological Reference Points for New England Groundfish. NEFSC Ref. Doc. 02-04. 417 p.

Northeast Fisheries Science Center (NEFSC) 2002b. Report of the 35th Northeast Regional Stock Assessment Workshop (35th SAW): SARC Consensus Summary of Assessments. NEFSC Reference Document 02-14. 259 p.

Northeast Fisheries Science Center (NEFSC) 2005. Report of the 41st Northeast Regional Stock Assessment Workshop (41st SAW): 41st SAW Assessmwnt Summary Report. NEFSC Reference Document 05-10. 36 p.

Parrack, M.L. 1986. A method of analyzing catches and abundance indices from a fishery. Int. Comm. Conserv. Atlantic Tunas, Coll. Vol. Sci. Pap. 24: 209-221.

Pauly, D. 1980. On the interrelationship between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. J. Cons. int. Explor. Mer 42: 116-124.

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

Prager, M.H. 1995. Users manual for ASPIC: a stock-production model incorporating covariates. SEFSC Miami Lab. Doc. MIA-92/93-55.

Ricker, W.E. 1954. Stock and recruitment. J. Fish. Res. Bd. Can. 11: 559-623.

Sipe, A.M., and M.E. Chittenden. 2001. A comparison of calcified structures for aging summer flounder, Paralichthys dentatus. Fish. Bull. 99:628-640.

Sissenwine, M.P., and J.G. Shepherd. 1987. An alternative perspective on recruitment overfishing and biological reference points. J. Cons. Int. Explor. Mer 40:67-75.

Smith, R.L, L.M. Dery, P.G. Scarlett, and A. Jearld, Jr. 1981. Proceedings of the summer flounder (Paralichthys dentatus) age and growth workshop, 20-21 May 1980, Northeast Fisheries Center, Woods Hole, Massachusetts. NOAA Tech. Memo. NMFS- F/NEC-11. 30 p.

Stock Assessment Workshop Southern Demersal Working Group (SDWG). 2004. Summer flounder assessment summary for 2004. 9 p.

Specker, J., R.R. Merson, C. Martinez, and B. Soffientino. 1999. Maturity status of female summer flounder and monkfish. URI/NOAA Cooperative Marine Education and Research Program (CMER) Final Report, Award Number NA67FE0385. 9 p.

Szedlmayer, S.T., and K.W. Able. 1992. Validation studies of daily increment formation for larval and juvenile summer flounder, Paralichthys dentatus. Can. J. Fish. Aquat. Sci. 49:1856-1862.

Terceiro, M. 1999. Stock assessment of summer flounder for 1999. Northeast Fisheries Science Center Reference Document 99-19, 178 p.

Terceiro, M. 2003. Stock assessment of summer flounder for 2003. Northeast Fisheries Science Center Reference Document 03-09, 179 p.

Terceiro, M. 2006. Summer flounder assessment and biological reference point update for 2006. http://www.nefsc.noaa.gov/nefsc/saw/2006FlukeReview/BRP2006_Review.pdf

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

Weber, A.M. MS 1984. Summer flounder in Great South Bay: survival of sublegals caught by hook-and-line and released. New York State Department of Environmental Conservation, Division of Marine Resources. Stony Brook, NY. 27 p.

Wigely, S., P.Hersey and J.E.Palmer. MS 2007. A description of the allocation procedure applied to the 1994 to present commercial landings data. Working paper in support of Terms of Reference A. GARM Data Review Meeting. http://www.nefsc.noaa.gov/GARMPublic/1.DataMeeting/A

Wilk, S.J., W. G. Smith, D.E. Ralph and J. Sibunka. 1980. The population structure of summer flounder between New York and Florida based on linear discriminant analysis. Trans. Am. Fish. Soc. 109:265-271.