A. WITCH FLOUNDER

TERMS OF REFERENCE

1. Characterize the commercial catch (landings and discards) through 2002.

2. Estimate fishing mortality, spawning stock biomass, and total stock biomass for 2002 and characterize the uncertainty of those estimates.

3. Evaluate and either update or re-estimate biological reference points as appropriate.

4. If stock projections are possible,
   a. provide short term projections (2003-2005) of stock status under various fishing mortality strategies and

   b. evaluate current and projected stock status against existing rebuilding or recovery schedules, as appropriate.

INTRODUCTION

The witch flounder (*Glyptocephalus cynoglossus, L.*) or grey sole is a deep water boreal flatfish occurring on both sides of the North Atlantic. In the Northwest Atlantic, witch flounder are distributed from Labrador to Georges Bank and in continental slope waters southward to Cape Hatteras, North Carolina. In U.S. waters, the species is commercially abundant in the Gulf of Maine-Georges Bank region [defined as Northeast Fisheries Science Center (NEFSC) Statistical Reporting Areas (SA) 511-515, 521-522, 525-526, and 561-562; Figure 1], and, in the absence of any stock structure information, is assumed to comprise a single stock unit. Prized as a table fish, witch flounder receives a high ex-vessel price relative to other flounders and represents an important by-catch component in the New England mixed species groundfish fishery. Annual landings during the period 1910-1982 averaged 3,000 metric tons (mt), ranging from 1,000 to 6,000 mt (Lange and Lux 1978, Burnett and Clark 1983). More recently, landings declined from a peak of 6,660 mt in 1984 to a low of 1,490 mt in 1990. Landings for 2002 were 3,186 mt.

Previous witch flounder stock assessments were conducted by Burnett and Clark (1983), Wigley and Mayo (1994) and Wigley et al. (1999). An assessment update was conducted for this stock in 2002 and reviewed at Groundfish Assessment Review Meeting (GARM; NEFSC 2002). The GARM assessment indicated average fishing mortality (ages 7-9, unweighted) increased from 0.21 in 1982 to 0.59 in 1985, declined to 0.24 in 1990, increased to 0.96 in 1996, then declined to 0.45 in 2001. Spawning stock biomass declined from 18,000 tons in 1982 to about 4,000 tons in 1995 and then increased to 11,368 mt in 2001. Since 1982, recruitment at age 3 has ranged from approximately 3 million fish (1984 year class) to 84 million fish (1997 year class) with a mean of 22 million fish. In 2001, the SSB was slightly above ½ SSBmsy (9,950 mt), the minimum stock size threshold, and fishing mortality (F= 0.45) was three times higher than Fmsy, the maximum fishing mortality threshold; thus, witch flounder was not overfished but overfishing was occurring in 2001.

This assessment of witch flounder in the Gulf of Maine - Georges Bank region and southward (USA Subareas 4, 5 and 6), presents a benchmark analytical assessment for the stock for the 1982-2002 period, estimates 2002 fishing mortality and spawning stock biomass for stock status, and provides short-term projections of median landings, discards and spawning stock biomass for various fishing mortality scenarios. This assessment provides estimates of discards from the shrimp fishery and large-mesh otter trawl fishery based upon analyses of sea sampling, commercial and research vessel survey data through 2002.

Witch flounder is managed under the New England Fishery Management Council's Multi-species Fisheries Management plan since 1987. A brief summary of groundfish management regulations affecting witch flounder is presented in Table A.1. Significant changes in regulations include increased minimum size in 1983 and 1987; increases in mesh size in 1982, 1983, 1994, 1999; effort reductions in 1996 and 2002; and implementation of closed areas in 1994 and 1998 (Figure 2). The western Gulf of Maine area closure, Cashes Ledge area closure and the seasonal rolling closures overlap the witch flounder distribution (Figures A.1 and A.2). Management regulations for the northern shrimp fishery also impact witch flounder (Table A.2); significant changes in the shrimp fishery include a monthly 10% by-catch limit which restricted the possession of groundfish to 10% by weight of shrimp in the mid-1980's to early 1990s; and the implementation of the Nordmore grate to exclude groundfish in 1992.

THE FISHERY

<u>Recreational Catches</u>
There is no recreational fishery for witch flounder.

<u>Commercial Landings</u>
USA commercial landings in 2002 totaled 3,186 mt, a 5% increase over 2001 (Table A.3); and 117% higher than in 1990, the lowest value since 1964 (Figure A.3). Canadian landings from the stock have been negligible (32 mt in 2001; Table A.3). Landings from the Grand Banks (NAFO Divisions 3LNO) during 1985 to 1990 are not included in this assessment. Canadian landings from the western Scotian Shelf (NAFO Division 4X) are not considered due to the fact that, until recently, witch flounder were reported as 'other flounders' by Canada, and cannot be separated from other flounder species. Furthermore, samples from the western Scotian Shelf indicate slower growth of witch flounder than in the Gulf of Maine, suggesting a different phenotypic stock.

The western Gulf of Maine (SA 513 and 514) and the central basin (SA 515) provide nearly a third of the USA witch flounder landings (Table A.4); landings from Georges Bank are confined to the deeper waters north of the South Channel (SA 521, 522; Table A.4). Otter trawl catches account for about 98% of witch flounder landings, with sink gillnets comprising the remainder (Table A.5). Catches are generally highest during March-July when witch flounder form dense pre-spawning aggregations (Burnett et al. 1992). The majority of witch flounder are landed in

Maine ports, primarily Portland, with lesser amounts landed in New Bedford and Gloucester, MA.

Although culling and grading practices vary by port, witch flounder have historically been landed as either 'small' or 'large'; however, three market categories ('peewee', 'medium', and 'jumbo') were added in some ports beginning in 1982 (Table A.6, Figure A.4). Since the early 1990s, the proportions of witch flounder landings from the peewee and small market categories have steadily increased. In 2002, witch flounder less than 45 cm ('peewee' and 'small' market categories) constituted 87% of total landings (Table A.6, Figure A.4). The current regulated minimum landing size for witch flounder is 36 cm (14 inches).

Sampling Intensity

Length frequency and age sampling data for witch flounder landings from the Gulf of Maine-Georges Bank region are summarized by quarter and market category in Table A.7 (because some ports do not cull into 'peewee' or 'jumbo' categories, NEFSC sampling protocols incorporate these categories into the 'small' and 'large' categories, respectively). Until 1982, sampling was minimal and sporadic. During 1982-1988, an average of 48 length frequency samples (approximately 100 fish per sample) was obtained annually over all market categories, representing 1 sample per 102 mt landed. In 1990, sampling requirements were adjusted to 1 sample per 50 mt to obtain more samples from the 'large' market category. However, samples for the 'large' market category have been difficult to obtain due to the sharp decrease in the landings of larger fish in recent years (Tables A.6 and A.7). Sampling intensity during 2001-2002 averaged 39 samples annually, representing 1 sample per 80 mt landed; nonetheless, even with this increased sampling intensity, inadequate numbers of samples were obtained for some market categories and quarter combinations. In 2002, of the 35 samples collected, 15 were small samples (43%), 10 were medium (29%) and 10 were large (29%). Compared with the 2002 market category landings distribution by weight (small 87%; medium: 10%; large: 3%), sampling in 2002 adequately approximated the market category distribution of landings on an annual basis. As in previous years, it was necessary to pool some quarters for some market categories. A summary of pooling procedures by year, market category and quarter is presented in Table A.8.

Commercial Landings at Age

Commercial age data for the years 1982 to 2002 were available for this assessment. Quarterly age-length keys (ALKs) were applied to corresponding commercial landings length frequency data by market category. Resulting estimates of annual age compositions (age 0 to 14+) are presented in Table A.9. No discernible changes in growth are evident during the 1982-2002 period; although landings mean weights and mean lengths at ages 6 to 8 declined in 1996-2002, this may be an artifact of poor sampling in recent years.

Discards

The Fisheries Observer Program (FOP), which began in 1989, has generated various levels of coverage for different fisheries. Prior to the FOP, NEFSC conducted sea sampling on an ad-hoc basis. The northern shrimp fishery, the small-mesh otter trawl fishery, and the large-mesh otter trawl fishery are three fisheries in which discarding of witch flounder occurs. In this assessment, discard estimates have been estimated for the shrimp fishery and the large-mesh otter trawl fishery.

Northern shrimp fishery
Since the 'shrimp season' spans a calendar year, in this report, the year in which most of the fishing occurred will be used to identify the entire season. For example, 1990 will refer to the shrimp season from December 1, 1989 to May 31, 1990. These estimation procedures were used in the 1994 assessment (Wigley and Mayo 1994), reviewed by the SAW 18 (NEFSC 1994), and extended through 1997 using the same methodology. The ratio of witch flounder discarded (kg) to days fished was calculated using FOP data for individual shrimp seasons, 1989-1997, by fishing zone. Since depth is an important factor influencing discards (Wigley MS 1994), discard ratios were calculated for each of three fishing zones (zone 1 = 0-3 miles from shore, zone 2 = 3 - 12 miles, and zone 3 = greater than 12 miles) in each season. For the most part, fishing zones are analogous to depth zones. Statistical testing of zonal discard rates indicated differences between fishing zones in most years. The zone-specific discard rates were weighted by the days fished in each zone to calculate a weighted mean discard rate for each season (Table A.10). To estimate witch flounder discard rates prior to the FOP, (i.e., 1982-1988), a simple linear regression was employed using 1989-1992 (years in which the Nordmore grate was not required) weighted mean discard rates and annual indices of witch flounder abundance. The NEFSC autumn bottom trawl survey index of age 3 fish was found to be the best predictor of annual discard rates (r^2 = 0.97, p = 0.0127; Figure A.5; Wigley MS 1994).

With no 1998-2002 FOP sampling in the northern shrimp fishery, an alternative method of survey filtering was explored to estimate witch flounder discard rates; however, due to insufficient length frequency data at small sizes, this method did not prove fruitful. As used for the years prior to the FOP, a simple linear regression using 1993-1997 (years in which the Nordmore grate was required) annual shrimp season discard rates and annual survey indices of autumn age 3 fish was employed (r^2 = 0.87, p = 0.0206). This five-point regression may not be as robust as the r^2 suggests, as four of the points are clustered (Figure A.5).

To obtain total weight of witch flounder discarded during a shrimp season, season discard rates (kg per day fished) were multiplied by the total number of days fished by the commercial fleet in each season (Table A.11). Estimated discard weight was then translated into discarded numbers at age by applying witch flounder sea-sampled discard length-frequencies expanded up to the total discard weight and then applying NEFSC spring bottom trawl survey ALKs. Detailed information on this method is given in Wigley (MS 1994).   For 1995-2002, days fished were estimated from the Vessel Trip Reports (VTR) using a stratification level of year, ton class, port group, month, and fishing zone. To derive the number of trips by fishing zone, the proportion of VTR trips by fishing zone was applied to the number of trips in the weighout database. Days fished per trip in each fishing zone were derived from the VTR data. Days fished per trip were then multiplied by the estimated number of trips for each fishing zone to derive estimated days fished by fishing zone, and then summed over year and fishing zone.

For the 1982-1997 time period, discard estimates of numbers at age and weight were derived on a shrimp season basis due to the limited number of length frequency samples in December. To adjust the shrimp fishery discard-at-age from a shrimp season basis to calendar year, the ratio of December days fished to the entire shrimp season days fished was used to apportion of the weight and numbers discarded into December and January-May categories. The December discards-at-age were shifted back one age, and then re-combined with the January-May matrix of the corresponding calendar year. The December discard weight was combined with January-May of the same calendar year. Mean lengths and mean weights at age in the re-combined catch at age were weighted by the numbers at age from each category.

Without 1998-2002 FOP sampling, discard length-frequency data were unavailable to partition the 1998-2002 estimated discard weight into numbers at length; thus discarded numbers at age were derived by apportioning discard weight by the average age composition (calendar year) of discards in 1993-1997 and then dividing by the average 1993-1997 discard mean weights at age. The average 1993-1997 mean weights at age from the FOP were consistent with trends in mean weights from the NEFSC survey during the 1998-2002 time period.

Witch flounder discards in the northern shrimp fishery ranged from a low 0.8 mt in 2002 to a high of 34 mt in 1988 and 1995 (Table A.11). Similarly, number of witch flounder discarded ranged from 40,000 fish discarded in 2002 to 1.8 million fish in 1994 (Table A.11). Estimates of age compositions of discarded witch flounder in the shrimp fishery are presented in Table A.12. Discarded witch flounder from the shrimp fishery range from age 0 to 6, with ages 1 to 3 most commonly discarded (Table A.12).

Large-mesh otter trawl fishery

Discard estimation from the large-mesh otter trawl fishery is confounded by the lack of FOP coverage prior to 1989, sparse coverage in the beginning of the program, and the recent implementation of year-around and seasonal area closures. As a result, three estimation scenarios were examined: 1) utilizing a survey filter method; 2) utilizing the at-sea observer data (Table A.13); and 3) utilizing the Vessel Trip Report data (Table A.14). The estimated discards (in weight and numbers) are presented in Table A.15. Each method is described below.

The method used in previous witch flounder assessments to estimate large-mesh otter trawl discards was based upon a method developed by Mayo et al. (1992) which utilizes survey and commercial catch at length data, commercial gear retention ogives, and information on culling practices. Research vessel length frequency data were filtered through commercial gear retention ogives corresponding to the predominant mesh size employed in the large-mesh fishery (130, 140, and 152 mm) and then through a culling practice ogive. Due to the sparse gear retention studies for witch flounder, mesh selection ogives were taken from Walsh et al. (1992) for American plaice. Given the high value and low abundance of this species, the culling practice of commercial fishermen was assumed to be nearly knife edged at the minimum landing size. A semi-annual ratio estimator of survey filtered 'kept' index to semi-annual numbers landed was used to expand the estimated 'discard' survey index to obtain numbers of fish discarded at length. The method used in this analysis differs from the method described by

Mayo et al (1992) which employs an expansion factor derived from a linear regression from the ratios of kept to landed at length. Semi-annual numbers of discard fish at length were apportioned to age using the corresponding season NEFSC ALK. Estimated numbers of discarded witch flounder in the large-mesh otter trawl fishery are presented in Table A.15. Results indicate that in recent years, numbers discarded at sea comprised as much as 54% of the witch flounder landed. The general pattern of discarding appears to be consistent with that expected given strong recruitment during 1979-1981 and the mid-1990's.

Given the distribution of juvenile witch flounder in the western Gulf of Maine and the recent implementation of year-around area closures and seasonal rolling area closures in the western Gulf of Maine, there was some concern regarding the application of the survey filter method to estimate discards in recent years. Since the commercial fishery does not have year-around access to the population estimated by the NEFSC survey, it may be inappropriate to use the survey filter method to estimate discards. For the 1989-2002 period, discard weight to kept weight ratios (D/K ratio) were calculated from FOP data on a semi-annual basis (Table A.13). Total discard weight was derived by multiplying the D/K ratio by the commercial landings. The number of sea sampled trips varied from no trips in the second half of 1992 to 83 trips in the second half of 2002. The D/K ratios ranged between 0.02 and 0.50. Given the limited number of trips, tows and available discard length frequencies, discards at age were derived only for the 1995-2002 time period (Tables A.15).

The Vessel Trip Report data were explored for information on discarding of witch flounder. Reporting of discard information in the logbooks is known to be incomplete. To eliminate problems associated with incomplete reporting, a subset of the VTR data was used. The VTR subset included only logbooks which reported discards of any species (Delong et al. 1997), assuming that operators who report discards of any species would reliably report witch flounder discards. This subset was used to estimate discard ratios (discard weight/kept weight) semi-annually for large-mesh otter trawl gear from 1994 to 2002. Limitations of this analysis are: 1) the dealer data used to expand discard rates to total discard weight do not contain information on mesh size, precluding partitioning of otter trawl fisheries into small and large mesh trips; 2) there is no area information on dealer data to isolate trips from the Gulf of Maine-Georges Bank region. From this analysis, results suggest that discard rates range between 4% and 9% (Table A.14). These estimates should be reviewed cautiously as not all fishermen report discards. Discarded numbers at age were estimated by expanding the FOP length frequencies and applying the survey age/length keys (Table A.15).

For estimates of total catch at age, discards from the large-mesh otter trawl fishery were derived using the survey filter method from 1982-1994 and using the FOP method for 1995-2002 (Table A.16).

Total Catch at Age

Total catch at age compositions (including commercial landings, discards from the northern shrimp fishery and the large-mesh otter trawl fishery) are presented in Table A.17 and Figure A.6. The age composition data reveal strong 1979-1981 year classes (Table A.17). The 1989 and 1993 year classes also appear to have been strong; however, these cohorts were heavily discarded in both the shrimp and large-mesh otter trawl fisheries (Tables A.12 and A.16). The poor 1984 year class is also evident as well as the truncated age-structure since the early 1990's.

Since witch flounder landings are highest during March-July, the average weights-at-age in the catch approximate mid-year weights. Mean weights at age at the beginning of the year (January 1; Table A.18) were derived from mid-year weights using procedures described by Rivard (1980).

STOCK ABUNDANCE AND BIOMASS INDICES

Commercial LPUE

Commercial catch rates (landings per unit effort, LPUE, expressed as landings in mt per day fished) were derived for vessel tonnage classes 2-4 [Class 2 consists of vessels 5 to 50 gross registered tons (GRT); Class 3, 51 to 150 GRT; and Class 4, 151 to 500 GRT]. These vessel classes account for greater than 95% of annual witch flounder otter trawl landings. LPUE indices for the Georges Bank-Gulf of Maine region were computed for: 1) all trips landing witch flounder, and 2) trips in which 40% or more of the total landings comprised witch flounder (Table A.19). These '40% trips' may represent effort that is 'directed' towards witch flounder, a species historically taken as by-catch.

For all trips landing witch flounder, increases in LPUE occurred in 1977-1978 for tonnage classes 2 and 3 and in 1982 for tonnage class 4, and remained high during the early 1980s; however, LPUE indices declined steadily for all tonnage classes from 1986 to 1990. Since the early 1990s, LPUE indices have steadily increased and are among the highest in the time series (Table A.19, Figure A.7a). Indices for 40% trips peaked in the early 1980's , then declined to a low in 1994, and have increased slightly in recent years (Table A.19, Figure A.7a). Effort (days fished) associated with all trips and 40% trips increased during the late 1970s and early 1980s, peaked during 1985-1988, and have generally declined since (Figure A.7b). While there is some evidence of increased directed effort in the early and mid 1980s [a period in which both witch flounder and American plaice were abundant and a small directed fishery emerged (Burnett and Clark 1983)], it is likely that LPUE indices derived for all trips landing witch flounder provide the best measure of relative abundance. In 1994 the NEFSC commercial data collection system changed from a voluntary to a mandatory system in which fishermen self-report fishing effort. Investigation is still on-going to determine if the time series of LPUE data can be extended (considered one series) or whether the post 1993 LPUE derived under the mandatory system constitutes a separate time series. Effort (days fished) for 1994 to 2002 may be underestimated in this report since effort is based upon preliminary VTR data, which do not represent 100% of the trips.

Research Vessel Survey Indices

The NEFSC has conducted annual research vessel stratified random bottom trawl surveys during autumn since 1963 and during spring since 1968. Details on survey sampling design and the use of survey data in stock assessments are given in Azarovitz (1981) and Clark (1981), respectively.

In September 2002, an offset in the trawl wraps was detected which may have effected the NEFSC bottom trawl surveys conducted from winter 2000 to the spring 2002. Extensive analyses of existing data sets and experimental studies were conducted to evaluate the offset issue (NEFSC 2002). These analyses were reviewed by a panel of experts and they concluded that no adjustments to the survey time series were justified (Groundfish Science Peer Review, 2003).

The Commonwealth of Massachusetts Division of Marine Fisheries (DMF) began an inshore trawl survey in 1978 which complements the NEFSC survey in coastal Massachusetts waters in that depths less than 27 meters (the lower depth limit sampled by the NEFSC offshore survey) are sampled (for details of this survey, see Howe et al. 1981). Additionally, the Northern Shrimp Technical Committee of the Atlantic States Marine Fisheries Commission (ASFMC) has conducted an annual northern shrimp survey during August in the Gulf of Maine since 1983, with catch data for witch flounder available from 1984 on (for details of the shrimp survey, see Northern Shrimp Technical Committee MS 1984). All three surveys provide useful information relative to trends in abundance, distribution, and recruitment of witch flounder in the Gulf of Maine-Georges Bank region. Strata utilized in the derivation of indices of relative abundance and biomass for witch flounder are as follows: NEFSC, offshore strata 22-30, 36-40 (Figure 3); Massachusetts DMF, regions 4 and 5; and northern shrimp, strata 1, 3, 6, and 8.

Witch flounder are generally distributed throughout the Gulf of Maine, along the Northern Edge and southern flank of Georges Bank, and southward along the continental shelf as far south as Cape Hatteras, NC (Figures A.1 and A.8). Juvenile witch flounder (< 25 cm) are distributed along the western Gulf of Maine, with a few in the canyon areas in the Mid-Atlantic region (Figures A.2a and A.2b). Concentrations of witch flounder along the western portion of the Gulf of Maine are observed in the ASMFC shrimp survey. Although this survey has limited spatial coverage (Figure A.9), most of the juvenile range is covered.

In response to a research recommendation from SARC 29, analyses were conducted to examine if the use of additional strata in the NEFSC bottom trawl survey might be appropriate. Burnett and Clark (1983) used NEFSC survey strata set 22,24,26-30, 33-40 in the first witch flounder assessment; however, Burnett (MS 1987) suggested that fish from strata 33, 34 and 35 exhibited different growth rates indicating these fish may be from a different stock inhabiting the western Scotian shelf. Based on this information, Wigley and Mayo (1994) revised the witch flounder survey strata set excluding 33, 34, and 35, and included strata 23 and 25 (Figure A.9). Following a method developed by Cadrin (2003), witch flounder catches for the entire autumn bottom trawl survey time series were examined by individual stratum. The stratified mean number per tow in each stratum was summed over the time period, and the percentage contribution of each stratum was calculated as well as the percentage of annual stratum sampling which produced no catch

(Table A.20). Results indicate that the current strata set (22-30, 36-40) accounts for approximately 93% of the survey catch and that only minor differences exist between the strata sets used in previous assessments. This analysis also indicated that stratum 6 contributed to the overall witch flounder catch. The stratified mean weight (kg) per tow was calculated for three strata sets: set 1 (22-30, 36-40); set 2 (22, 24, 26-30, 36-40); and set 3 (6, 22, 24, 26-30, 36-40). The trends of these biomass indices (and their variance) are indistinguishable (Figures A.10a and A.10b). The inclusion of stratum 6 is not justified due to its geographical discontinuity with the core strata. Since no additional strata were identified as contributing to the total catch, or improved the precision of the estimates of mean weight per tow, the strata set 22-30, 36-40 will continue to be used.

Research vessel survey indices of abundance, biomass, and mean length for NEFSC surveys, Mass. DMF surveys, and ASMFC shrimp surveys are presented in Tables A.21-A.23 and Figures A.11-A.16, respectively. A summary of available age data from NEFSC surveys is given in Table A.24; survey age samples collected during 1976 to 1979 have not been aged. Too few age samples are collected during DMF surveys to reliably characterize the age composition of witch flounder in the inshore areas, and no age samples are collected on ASMFC surveys. Age-specific relative abundance indices from NEFSC spring and autumn surveys 1980-2002, and preliminary spring 2003 are presented in Table A.25, Figures A.17 and A.18. Mean length and mean weights at age from the NEFSC spring and autumn surveys area given in Tables A.26 and A.27 and Figures A.19 - A.21.

While NEFSC spring survey indices tend to be more variable due to the pre-spawning aggregations of witch flounder, spring and autumn indices generally display similar trends. Abundance and biomass remained fairly stable from 1963 until the late 1970s (Table A.21, Figures A.11 and A.12); autumn indices declined during the early and mid 1980s, reaching record low levels in 1987. Abundance sharply increased in 1993, due to a large age 0 index (Table A.25, Figure A.12) and has continued to increased to near record high levels in 2002. During the same time, mean length declined (Figures A.15 and A.16). The age structure has been truncated since the late 1980's (Figures A.17 and A.18).

Length frequency data from the ASMFC shrimp survey suggest that incoming year classes can be identified prior to their appearance in the NEFSC surveys. Thus, the ASMFC survey appears to be more useful in providing a pre-recruit index than in characterizing the population as a whole (Table A.23). The ASMFC survey data indicate improved recruitment in recent years, corresponding to age 1 fish, during 1991-1994, 1997, and 1999. Significant numbers of small fish were also observed in the NEFSC autumn survey during the same years.

Mean lengths at age from NEFSC spring and autumn surveys are presented in Table A.26 and for ages 4 to 8 in Figures A.19a and A.19b. Mean lengths at age for ages 5 to 7 appear to have increased approximately 3-5 cm from 1980 to the late 1980's, and then declined (Figures A.19a and A.19b); however, Von Bertalanffy growth analyses detected no significant changes in resulting growth parameters over the time period.

NEFSC spring and autumn survey mean weights at age are given in Table A.27 and Figures A.20 and A.21. Survey mean weights are variable, however, similar declines in mean weights for ages 6-9 were observed during the mid-1990s to 2002 in both the commercial landings and spring and autumn surveys.

MATURITY

Witch flounder maturity observations have been collected on the NEFSC research bottom trawl surveys since 1977. The NEFSC spring surveys were used for maturity analyses as these surveys occur closest to and prior to spawning (Halliday 1987). In the previous witch flounder assessment, probit analyses (SAS 1985) of maturity at age data revealed that there have been six maturity stanzas over the assessment period (GARM NEFSC 2002). The proportion at which 50% of the fish are mature at age (A_{50}) was significantly different for the time periods 1980-1982, 1983-1984, 1985-1990, 1991-1993, 1994-1999, and 2000-2002. Due to small sample sizes, it was necessary to pool individual years, however, individual years were examined, and then pooled into time blocks. Trends in female A_{50} and L_{50} were similar, progressively decreasing from 1980-1982 to 1985-1990, then increasing in 1991-1993, then declining in 1994-1999 and increasing in 2000-2002 similar to 1983-1984 levels. The maturity stanzas used revealed sharp changes in proportion mature, uncharacteristic of the assumed gradual biological process. The maturity stanzas also revealed, in a few instances, biologically infeasible outcomes, i.e. over the life span of a cohort, the proportion mature at age would decrease. Given these issues, a method which has been applied to Georges Bank cod (L. O'Brien, NEFSC, pers. comm.) was employed to minimize the abrupt changes yet still capture the changing trends in maturity over time. This method used logistic regression and a five-year moving time block to estimate annual maturity ogives. For example, the proportion mature in 1982 was estimated using NEFSC spring maturity data from 1980, 1981, 1982, 1983 and 1984. Likewise, the 1983 maturity ogive used maturity data from 1981 to 1985. Annual maturity ogives were derived for 1982 to 2001 using 1980 - 2003 data. The annual 2002 maturity ogive was assumed to be equal to the 2001 ogive (Table A.28, Figure A.22). In addition to the annual maturity ogives, a single ogive using maturity data from the entire time series was also calculated (Table A.28). It was concluded that the moving time block method was appropriate for use in the VPA.

Stratified mean weight per tow of mature (spawning stock) witch flounder was calculated for spring NEFSC research vessel surveys (Table A.29, Figure A.23) using the six maturity stanzas. This analysis will be updated to incorporate the moving time-block maturity estimates in the next assessment update. The spawning stock biomass indices closely track total biomass indices except in most recent years, indicating a larger proportion of immature fish in the population.

MORTALITY

Natural Mortality

Burnett (MS 1987) estimated instantaneous natural mortality (M) to be 0.16 from a regression of survey-derived instantaneous total mortality (Z) estimates on commercial fishing effort. Halliday (1973) used a value of M = 0.15 for females and M = 0.2 for males in an assessment of Scotian Shelf witch flounder. In the present study, virtual population analyses, yield per recruit and spawning stock biomass per recruit analyses were performed assuming M = 0.15.

Total Mortality

Estimates of instantaneous total mortality (Z) were computed from NEFSC spring and autumn research vessel bottom trawl survey catch per tow at age data by combining cohorts over the following time periods: 1982-1985, 1986-1989,1990-1993, 1994-1997 and 1997-2001. Given the variability in age at full recruitment to the sampling gear observed during the survey time series (Table A.30), estimates were derived for each time period and each season by taking the natural logarithm of the ratio of pooled age 7+ to pooled 8+. For example, the estimate of Z for 1982-1985 was computed as:

Spring: ln (sum age 7+ for 1982-1985 / sum age 8+ 1983-1986)
Autumn: ln (sum age 6+ for 1981-1984 / sum age 7+ 1982-1985).

To evaluate Z over identical year classes within each of the survey series, different age groups were used in the spring and autumn.

Total mortality estimates from the two survey series exhibited similar trends, although autumn estimates were generally lower than those in the spring (Table A.30 and Figure A.24a). With no objective basis to select one survey series over another, total mortality was calculated by taking the geometric mean of the spring and autumn estimates during each time period. Total mortality ranged between 0.34 and 0.71 over the time series (Table A.30). Additionally, annual estimates of total mortality were calculated, and smoothed with a three year moving average (Figure A.24b).

ESTIMATION OF FISHING MORTALITY RATES AND STOCK SIZE

Virtual Population Analysis and Calibration

The ADAPT calibration method (Parrack 1986, Gavaris 1988, Conser and Powers 1990) was applied to estimate abundance at age in 2003 using catch-at-age estimates (i.e., landings plus discards from the shrimp and large-mesh otter trawl fishery; Table A.17). Estimates of stock sizes, their associated statistics, and F in the terminal year are summarized in the Table A.31.

New VPA software is now available in the NOAA Fisheries Toolbox. To bridge the transition between the software used in the last assessment update (FACT 1.5) and the current software, NFTv2.0.11, the accepted 2002 VPA (NEFSC 2002) formulation and input data was re-run using the NFTv2.0.11 software. The summary statistics of the two VPAs (RUN 61-f) reveal only slight changes in stock size estimates and fishing mortality (Table A.31), and these minor changes are attributed to the use of the exact catch equation and other improvements in precision.

An initial formulation (RUN 100) based upon the 2002 VPA was performed to estimate 2003 stock sizes for ages 4 to 10 (Table A.31) using a catch-at-age matrix including ages 3-11+ and NEFSC spring and autumn abundance indices for ages 3 to 11+ as tuning indices. All indices were given equal weighting. Autumn survey indices were lagged forward one year and one age to calibrate with beginning year population sizes of the subsequent year. A flat-top partial recruitment (PR) pattern was assumed, with full fishing mortality on ages 7 and older. The F on ages 10 and 11+ in the terminal year was estimated as the average of F on ages 7 through 9. The F on ages 10 and 11+ in all years prior to the terminal year was derived from weighted estimates of Z for ages 7 through 9. Instantaneous rate of natural mortality (M) was assumed to be 0.15. Spawning stock biomass (SSB) was calculated at time of spawning (March) and mean weight at age calculated by the Rivard method (Table A.18).

The results of the initial run indicated that coefficients of variation (CV) for estimated ages ranged between 29% and 44% and the CVs for survey catchability coefficients (q) were consistent, ranging from 11% to 27%.

Two alternative formulations included: 1) using a total catch at age in which large-mesh otter trawl discards were estimated using the survey filter method for 1982-1994 and Fisheries Observer Program data for 1995 to 2002 [RUN 200]; and 2) estimating age 3 stock size using survey tuning indices [RUN 201]. Results from these alternative formulations provided estimates of stock size, F and spawning stock biomass consistent with the base run [RUN 100]. RUN 201 stock size for age 3 was poorly estimated (CV = 63%). Based on these runs, the partial recruitment pattern indicated that age 7 was not fully recruited. An alternative formulation (RUN 300) was conducted using a partial recruitment vector where the fully recruited age was increased from 7 to 8. Assuming full recruitment at age 8, the F on ages 10 and 11+ in the terminal year was estimated as the average of F on ages 8 and 9. The F on ages 10 and 11+ in all years prior to the terminal year was derived from weighted estimates of Z for ages 8 through 9. This partial recruitment pattern is consistent with recent mesh regulation changes.

The final formulation (RUN 301-f) included a 3 to 11+ catch at age with large-mesh otter trawl discards estimated using both the survey filter method and FOP data; an updated partial recruitment vector reflecting current management regulations was derived from the 1999-2002 F pattern taken from a penultimate calibration run; annual maturity ogives estimated by the five year moving time block with the 2002 maturity vector assumed to be equal to 2001. Ages 1 and 2 were deleted from the catch at age, this allowed recruitment in 2003 to be estimated using the the geometric mean; there is no difference between 1-11+ vs 3-11+ on VPA results for fishing mortality and spawning stock biomass. Based on the final formulation, two sensitivity analyses were conducted to evaluate the selection of tuning indices. The VPA was tuned with only

NEFSC spring survey indices and then tuned with only NEFSC autumn survey indices (Table A.31). Estimates of F and SSB from analyses using a single tuning series bounded the F and SSB estimated using both spring and autumn tuning indices. Using only the spring tuning series (RUN 301-f-spr), F was slightly higher (F= 0.43) and SSB is slightly lower (15,798 mt) then the final run (RUN 301-f). Conversely, using only the autumn tuning indices (RUN 301-f-aut), F is slightly lower (F = 0.39) and SSB is slightly higher (21,569 mt; Table A.31) then the final run (RUN 301-f) .

VPA Estimates of Fishing Mortality, Spawning Stock Biomass and Recruitment

The VPA results, including estimates of F, stock size and spawning stock biomass at age are given in Tables A.32. The mean residual for the VPA calibration was 0.791 and the CV on age 3-10 stock sizes ranged from 31% to 64% while the CVs on the estimates of survey catchabilities were between 13% and 26%. The normalized survey indices and standardized residuals are presented in Figures A.25 And A.26.

The VPA indicates that fishing mortality (ages 8-9, unweighted) increased from 0.26 in 1982 to 0.67 in 1985, declined to 0.22 in 1992, increased to 1.13 in 1996, then declined to 0.41 in 2002 (Table A.33 and Figure A.27). Spawning stock biomass declined from 16,897 mt in 1982 to about 3,800 mt in 1996. With recent increases in recruitment and declines in fishing mortality, SSB has increased to 18,296 mt in 2002 (Table A.33 and Figure A.28). Since 1982 recruitment of age 3 has ranged from approximately 3 million fish (1984 year class) to 67.6 million fish (1997 year class; Table A.33 and Figure A.28). Over the 1982-2002 period, average recruitment of age 3 fish (the 1979 - 2000 year classes) was 19.6 million (the geometric mean equaled 14.4 million fish). The 1995-1999 year classes appear to be above average, and the 1997 year class is the largest in the VPA time series (Table A.33 and Figure A.28) .

The relationship between spawning stock biomass and recruitment (age 3) is presented in Figure A.29. The negative stock-recruitment relationship observed in previous assessments continues with the addition of the 2000 year class.

Precision of F and SSB

The uncertainty associated with the estimates of stock size and fishing mortality from the final VPA was evaluated using a bootstrap procedure (Efron 1982). One thousand bootstrap iterations were performed to derive standard errors, coefficients of variation (CVs) and bias estimates for the stock size estimates at the start of 2003, the catchability estimates (q) of the abundance indices used in calibrating the VPA, and the 2002 fully recruited fishing mortality rate (age 8+). Frequency distributions of the 2002 mean fishing mortality and spawning stock biomass bootstrap estimates were generated and cumulative probability curves produced (Figures A.30 and A.31).

Bootstrap results suggest that the estimates of 2003 abundance had CVs between 32% to 84%, 24% for 2002 $F_{8\text{-}9}$ and 15% for 2002 spawning stock biomass. There is an 80% probability that the 2002 F (0.41) lies between 0.31 and 0.56 (Figure A.30), and the 2002 SSB (18,296 mt) lies between 15,603 mt and 22,969 mt (Figure A.31).

Retrospective Analyses

A retrospective analysis was conducted on the final VPA (Run 301-f) from 2002 to 1992 by sequentially removing the terminal year of the data to evaluate internal consistency of the current ADAPT formulation with respect to terminal estimates of F, SSB, and recruits at age 3 for the seven years prior to the current assessment. Results indicate that average F was underestimated (Figure A.32a) and spawning stock biomass was consistently overestimated (Figure A.32b). The retrospective analysis indicated that the number of age 3 recruits were generally overestimated, and the 1995-1997 year classes were considerably overestimated (Figure A.32c).

Statistical Catch-at-age model

A statistical catch-at-age analysis was conducted for the witch flounder stock. An age-structured forward-projection model (a.k.a., age-structured production model) was fit to fishery and survey data during 1937-2002. This model provided an alternative long-term perspective on resource dynamics in comparison to VPA-based analyses that were limited to the period 1982-2002. Age-structured population dynamics of witch flounder were described using forward-projection methods for statistical catch-at-age analyses (Fournier and Archibald 1982, Methot 1990, Ianelli and Fournier 1998, Quinn and Deriso 1999). Models were fit to data with the AD Model builder software for nonlinear optimization (Otter Research 2001).

Six alternative statistical catch-at-age models were developed and fit. Brodziak and Wigley (2003 ms) contains a complete description of the basic model and input data. Common features of the six models were:

- Natural mortality was M=0.15 for all age classes.
- Catch scenario 2 was used (same catch as used in the VPA).
- Fishery selectivity was estimated for historic (1937-1993) and current (1994-present) time periods.
- NEFSC spring and fall survey biomass and numbers at age data were used.
- Emphasis values for likelihood components were:Recruitment λ_1=10, Fishery age composition λ_2=1, NEFSC Fall survey age composition λ_3=1, NEFSC Fall survey biomass index λ_4=100, NEFSC Spring survey age composition λ_5=1, NEFSC Spring survey biomass index λ_6=100, Catch biomass λ_7=100, Fishing mortality λ_8=1, Fishing mortality penalty λ_9=1

The primary differences among the six alternative models were:

1. Dome-shaped selectivity possible for fishery, spring, and fall surveys; time frame is 1937-2002.
2. Flat-topped selectivity for fishery, spring, and fall surveys; time frame is 1937-2002.
3. Dome-shaped selectivity possible for fishery, spring, and fall surveys; time frame is 1963-2002.
4. Flat-topped selectivity for fishery, spring, and fall surveys; time frame is 1963-2002.
5. Flat-topped selectivity for fishery and spring survey; Dome-shaped selectivity possible for fall survey; time frame is 1963-2002.

6. Flat-topped selectivity for fishery and spring survey; Dome-shaped selectivity possible for fall survey; time frame is 1937-2002.

Models 1, 2, and 6 were considered to be the primary models, while models 3, 4, and 5 provided sensitivity analyses to the choice of time frame. The Northern Demersal Working Group (WG) reviewed the model diagnostics. In general, the selectivity patterns of models that allowed dome-shaped fishery or survey selectivity appeared to be too sharply domed to be biologically plausible. In contrast, models with the assumption of flat-topped selectivity provided a poorer fit to the data, as measured by the root-mean squared errors for the NEFSC fall and spring survey biomass index and the catch biomass fits to the data. The WG chose to reduce the emphasis on the NEFSC fall and spring survey biomass index and the catch biomass likelihood components to 10 down from 100. This choice alleviated the problem of implausible selectivity patterns in the fishery and the survey. As a result, the WG concluded that Model 1 with reduced emphasis values was the best alternative of the statistical catch-at-age analyses (SCAA). Model results are reported to confirm the basic trends of VPA-based results and show the likely effect of extending the assessment time horizon back to 1937.

Model results showed that current fishery selectivity at age was estimated to be lower at ages 1-6 than historic selectivity (Figure A.33). This was consistent with increases in fishery mesh size and changes in discarding practices (e.g., shrimp fishery) that occurred around 1994. The resulting catch biomass predictions generally matched observed catch biomasses (Figure A.34) with some moderate deviations in the early 1980s.

Model results showed that NEFSC fall survey selectivity was dome-shaped with a peak at age-5 (Figure A.35). The NEFSC spring survey selectivity was flat-topped with full selection occurring at roughly age-7. The resulting predicted NEFSC fall and spring survey indices generally matched the trends in observed indices (Figures A.36 and A.37). Both surveys indicate a longterm decline in biomass from the 1970s through the early 1990s. Biomass increases in the late-1990s differed moderately between the fall and spring surveys

There was general agreement between VPA and SCAA results during 1982-2002. Spawning biomass estimates were very similar during 1989-1999 (Figure A.38, SCAA estimate of 10.5 kt in 2002). The VPA indicates a smaller decrease in spawning biomass during 1982-1988 and a greater increase during 2000-2002. Fishing mortality estimates were also similar (Figure A.39, SCAA estimate of 0.48 in 2002). Both VPA and SCAA estimates increased to roughly the mid-1990s and then declined. Recruitment estimates also exhibited similar patterns (Figure A.40, SCAA estimate of 14.1 million age-1 fish in 2002), although the VPA indicated larger increases in recruitments during the late-1990s. Despite differences in model configuration and estimation approach, the SCAA generally confirmed point estimates and trends in the VPA results.

BIOLOGICAL REFERENCE POINTS

Yield-per-recruit (Y/R) and spawning stock biomass per recruit (SSB/R) analyses were performed using the Thompson and Bell (1934) method for witch flounder ages 3 to 20. Input vectors for partial recruitment, maturation at age and mean weights at age were all updated since the last assessment. Mean weights at age used in the Y/R analyses were computed as an arithmetic average of catch mean weights at age (Table A.17) over the 1999-2002 period. Mean weights at age for use in the SSB/R analyses were derived by applying the length-weight relationship for witch flounder to predicted lengths at age from von Bertalanffy growth curve analyses of NEFSC survey data from 1980-2002. The maturation ogive from the entire time series (1980-2003) was also used (Table A.28). Given the changes in regulated mesh size in 1999, the exploitation pattern used in the yield and SSB per recruit analyses and short-term projections was computed from the 1999-2002 VPA results. Geometric mean F at age was computed for the 1999-2002 period and divided by the geometric mean of the fully recruited annual Fs to derive the partial recruitment vector. The final exploitation pattern was smoothed, applying full exploitation on ages 8 and older, viz.

Age 3	Age 4	Age 5	Age 6	Age 7	Ages 8+
0.0036	0.0229	0.0703	0.1931	0.5282	1.000

The input data and results for the Y/R and SSB/R analyses are given in Table A.34 and Figure A.41. The reference points were $F_{0.1}$ = 0.196, F_{max} = 0.545, and $F_{40\%}$ = 0.230.

The biological reference points were updated by applying the approach used to estimate MSY proxies for witch flounder (NEFSC 2002). Fmsy is approximated as F40% (0.23), the SSBmsy proxy is 25, 248 mt, the product of 40%MSP (1.2882 kg spawning biomass) and average long-term recruitment (19.6 million). The MSY proxy is 4,375 mt, the product of yield per recruit at F40% (0.2232 kg) and average recruitment.

.

In 2002, spawning stock biomass was slightly greater than ½ SSBmsy (12,624 mt), the minimum stock size threshold, and fishing mortality in 2002 was nearly double Fmsy, the maximum fishing mortality threshold; therefore, witch flounder was not overfished but overfishing was occurring in 2002 (Figure A.42).

To evaluate the effects of simultaneous changes in the three input vectors described above (i.e. partial recruitment, maturation and mean weights) on F40%, Y/R, SSB/R and the SSBmsy proxy, a decomposition analysis (P.Rago, NEFSC, pers. comm.) was conducted. This analysis is analogous to decomposing a sum of squares in an analysis of variance (decomposing the total resulting difference into its components).

For F40%, Y/R and SSB/R:
Total effect = effect of vector 1 + effect of vector 2 + effect of vector 3 + interaction terms.

For SSB_{MSY}:
Total effect = SSB/R effect + Recruit effect + interaction term.

The effect is the difference between the former YPR estimate and the current YPR estimate, for F40%, Y/R, SSB/R and for SSB_{MSY}.

To accomplish this, the former YPR analysis (Run 0) was re-run using ages 3-20 to coincide with the ages used in the current YPR analysis (Run 1). Then, YPR analyses were conducted where each former vector was replaced with a current vector (Runs 2 through 7), until all vectors were replaced with current vectors (Run 8). The resultant F, Y/R and SSB/R at 40%MSP from each run (Runs 0 to 8) are reported in Table A.35. The total effect of changing all three vectors at once equals Run 1 - Run 8.

Results of the decomposition analysis (Table A.35) indicate that changes in F40% were effected most by new partial recruitment vector. Changes in Y/R resulted from the interaction of all three new input vectors while changes in the SSB/R resulted from the interaction between the mean weights and maturity vectors. Changes in SSBmsy were effected most by changes in new mean age 3 recruitment.

SHORT-TERM PROJECTIONS FOR 2004 AND 2005

Short-term stochastic projections were performed to estimate landings, discards and SSB during 2003-2005 under various F scenarios using bootstrapped VPA calibrated stock sizes in 2002 The partial recruitment, maturity ogive, and mean weights at age were the same as described in the yield and SSB per recruit section (Table A.36). Recruitment (age 3) in 2003-2005 was derived by re-sampling the cumulative density function based on the empirical observations during 1982-2002 (1979-2000 year classes). Fishing mortality was apportioned among landings and discards based on the proportion observed landed at age during 1999-2002. The proportion of F and M which occurs before spawning equals 0.1667 (March 1); M was assumed to be 0.15. Spawning stock biomass in 2002 was estimated to be 18,296 mt. The F scenarios are: status quo F_{2003} = 0.41, Fmsy = 0.230, 75% of Fmsy = 0.17 and landings_{2003} = landings_{2002} (F= 0.199). Fishing at the status quo F (0.41) or at the target (Fmsy = 0.23) in 2003 - 2005 is expected to allow biomass to increase above SSBmsy and initiate rebuilding of the age structure (Table A.36). Comparison of the current age structure and the age structure under MSY conditions are given in Figure A.43.

CONCLUSIONS

Based on the ADAPT VPA, the witch flounder stock was not overfished, but overfishing was occurring in 2002. Fully recruited fishing mortality in 2002 was 0.41, nearly double Fmsy (0.23), and spawning stock biomass was estimated to be 18,296 mt in 2002, 72% of SSBmsy (25,248 mt) . Recent year classes appear to be above average. Although the spawning stock biomass has increased, the age structure still remains truncated. Fishing mortality should be reduced to Fmsy or below to allow the age structure to rebuild.

WORKING GROUP DISCUSSION

The Working Group noted the truncated age structure in the landings during the 1990s, and concluded that the 11+ group was appropriate for this species. The Working Group discussed the survey filter method and its potential to overestimate discards when closed area exists. The Working Group concurred that the survey filter method should be used only when Fisheries Observer Program data are not available or insufficient to characterize discards. The Working Group accepted the large-mesh otter trawl discards which had been estimated using both the survey filter method for the 1982-1994 period and the FOP data for the 1995-2002 period.

The maturity analyses and limitations of using multiple maturity stanzas was also discussed. The Working Group examined the annual estimates of A50 over time relative to the six stanzas and agreed that, while time trends in A50 were evident, the multi-year moving time block method used to estimate annual maturity ogives was appropriate.

The Working Group pointed out that very few witch flounder are caught during the NEFSC bottom trawl surveys. In most years, the stratified mean number per tow of witch flounder is less than five fish. During the late 1980's and early 1990's, the abundance of witch flounder may have gone below detectable levels with one or less than one fish per tow.

The Working Group pointed out that the recent, above-average year classes may be poorly determined, and based on the retrospective pattern for recruitment, these year classes may be overestimated.

SARC DISCUSSION

The effect of low sampling intensity of witch flounder in commercial landings was discussed. It was noted that in recent years, the sampling ratio has decreased especially in the small market category (87% percent of commercial landings), and it was recommended that the commercial sampling be allocated appropriately to the landings of each market category.

The SARC noted that there has been a recent increase in the proportion of smaller witch flounder in the NMFS survey, as well as the truncation of older age classes. The truncation of the age structure in the survey is consistent with high fishing mortality in the 1990s. Since full recruitment is estimated to occur at age eight, there is concern that age truncation could have a serious effect on the future reproductive potential. However, since the 1997 and 1998 above average recruitment events, the age structure is starting to expand compared to the early 1990s when fishing mortality was high and recruitment of the 1983 and 1984 year classes was very poor.

The SARC discussed the difference between survey selectivity estimates in the VPA and SCAA models. In the SCAA, there appears to be a flat-top pattern for the spring and a dome-shaped pattern for the autumn NEFSC survey. The VPA model indicates that the two surveys have similar selectivity. The SARC noted the recent decrease in mean length per tow of witch flounder, which is more apparent in the autumn compared to the spring survey. The greater inter-annual variability of mean length per tow in the autumn survey may be magnifying the discrepancy between the two models. Alternatively, the differences in selectivity and mean weights between the surveys could be a manifestation of seasonal pre-spawning aggregations of witch flounder that differentially affect the availability of older or younger witch flounder to the surveys.

The SARC discussed the decline in the mean weights at age, since declines in mean weight are counter-intuitive compared to the usual response of stocks to overfishing. There were three possible hypotheses: 1) a fishery effect; 2) a density dependent effect or 3) an environmental effect. Fishery effects on mean weights at age can occur when a fishery tends to catch larger fish of a cohort as often occurs for partially-recruited age classes. The declining mean weights at age occurred in fully-recruited year classes, however. Changes in mean weight at age are also commonly observed as a response to year class strength, a density-dependent effect. On the other hand, the year class strength was low for the older year classes where declining mean weights were observed. The other possibility, suggested by the SARC, is that environmental effects may have caused changes in large fish distribution or in growth rates which may or may not reverse as the stock size increases.

Also noted was the peculiar stock-recruitment pattern for witch flounder, where the strongest year classes were spawned when SSB was lowest, and vice versa. The SARC thought that further investigation was needed to evaluate the productivity of witch flounder and the value of biomass targets derived from assumptions about the estimated relationship between spawning biomass and recruitment. It was recommended that a longer time series of data from the SCAA results be explored as well as examining the relationship with trends in abundance of primary predators identified in the food habits data base or other factors to identify possible causes for the observed recruitment pattern. SARC analyses of NEFSC food habits data revealed that witch flounder occurred as prey items in 64 predators (13 species) during the 1978 to 2000 period.

There were several methodological issues raised by the SARC. The SARC commented that estimation uncertainty for input parameters were not included in the YPR analysis. Concerning the ability to compare and evaluate different model formulations and methods, the SARC recommended using an overall statistic (e.g. AIC statistic) for this purpose.

For the accepted VPA formulation, the SARC noted that the CVs are within acceptable ranges and residuals do not show strong patterns that would indicate the model's lack of fit to the data, although in older ages, there may be some positive bias in the earlier years of the time series. It was noted that the uncertainty of the assumptions associated with the model were not addressed since the 80% CI estimates only reflect uncertainty of the model fit and does not incorporate all sources of uncertainty. The SARC also discussed the justification for using F40% as the basis for estimating a SSB proxy as a substitute for Bmsy. It was concluded that using F40% was an acceptable parameter to use for a slow-growing, late-maturing, flatfish species. Changes in the SSBmsy proxy value were attributed mainly to the addition of new recruitment data that included the strong 1995 - 2000 year classes occurring during 1998 to 2002.

The SARC considered an alternative statistical catch at age model (SCAA) for comparison with the VPA results. The SCAA approach can potentially account for uncertainty in the catch and incorporate information when the full catch-at-age data are not available (e.g. historical landings). The SARC agreed to accept the VPA assessment, although the SCAA model is under development and is giving comparable estimates. The SARC commented that the projections using the terminal year estimates of numbers at age may be optimistic given the retrospective patterns of the VPA. Uncertainties in the discard estimates may also be contributing to the observed retrospective pattern. The SCAA model avoids this problem because it accounts for errors in the catch.

The SARC also noted the landings were comprised of small fish, the age-structure of the population was truncated, and the changes in growth and maturity were occurring. There is concern that average recruitment from the VPA time series may overestimate average recruitment over all stock sizes in the projections. The SARC was also concerned about the reliability of projections since the SSB includes young spawning fish and that the retrospective patterns tend to be optimistic. It was noted that the VPA estimated higher stock biomass in recent years than those estimated by the SCAA model.

SOURCES OF UNCERTAINTY

- The research bottom trawl survey catches very few witch flounder; in most years, the stratified mean number per tow of witch flounder is less than 5 fish. Abundance of witch flounder in the late 1980s and early 1990's may have gone below levels that provide reliable estimates of trends in abundance and biomass.

- Low sampling intensity of commercial length samples across market category and quarter, especially seen in the recent decreased ratio of small market category sampling, results in imprecise mean weights at age and estimates of numbers at age.

- The VPA calibration may be confounded because survey-based estimates of discards use the same information as that used as tuning indices. Survey information was used to estimate discards for the large-mesh otter trawl fishery during 1982 to 1994, as a substitute for the lack of FOP data prior to 1989 and sparse FOP coverage through 1994.

- Retrospective patterns suggest that 2002 SSB may be overestimated (i.e. future assessments may provide lower estimates of 2002 SSB) and fishing mortality may be underestimated (i.e. future assessments may provide higher estimates of F).

- Various factors including selectivity ogives, mean length of discards, and sampling frequency, introduce uncertainties in the VPA that are not appropriately treated because the VPA assumes that catches are known without error.

RESEARCH RECOMMENDATIONS

- Continue to develop alternative models to the VPA, focusing on those that incorporate sampling error and uncertainties in input parameters. While the statistical catch-at-age model is a useful approach, it is still at the developmental stage; further work is needed to examine the sensitivity of the model's weighting factors.

- Investigate the sensitivity of SSB estimates to the number of years used to calculate annual proportion mature at age using a multiple year time block.

- Explore the sensitivity of the assessment models to discard at age estimates, especially with respect to retrospective patterns and other diagnostics.

- Explore the usefulness of the Maine Department of Marine Resources inshore survey for estimating trends in relative abundance and biomass, and for use as assessment tuning indices.

- Improve the biological sampling of all market categories. Sampling should be proportion to landings.

LITERATURE CITED

Azarovitz, T.R. 1981. A brief historical review of the Woods Hole Laboratory trawl survey time series. *In:* Doubleday, W.G. and D. Rivard (eds.), Bottom Trawl Surveys, p. 62-67. Can. Spec. Publ. Fish. Aquat. Sci. 58.

Brodziak, J., and S. Wigley. 2003 manuscript. Statistical catch-at-age analysis of the witch flounder stock, 1937-2002. Northern Demersal Working Group Meeting May 12-16 2003, Working Paper W3, NEFSC, Woods Hole, MA 02543.

Burnett, J.M. MS 1987. The population biology of the witch flounder, *Glyptocephalus cynoglossus* (L.), in the Gulf of Maine-Georges Bank region. M.Sc. Thesis, Department of Wildlife and Fisheries Biology, University of Massachusetts, Amherst, MA, 116 p.

Burnett, J. and S.H. Clark. 1983. Status of witch flounder in the Gulf of Maine - 1983. NMFS/NEFC, Woods Hole Laboratory Ref. Doc. No. 83-36, 31 p.

Burnett, J., M.R. Ross, and S.H. Clark. 1992. Several biological aspects of the witch flounder (*Glyptocephalus cynoglossus* (L.)) in the Gulf of Maine-Georges Bank region. J. Northw. Atl. Fish. Sci. 12: 15-25.

Cadrin, S.X. 2003. Stock assessment of yellowtail flounder in the southern New England - Mid-Atlantic area (SAW 36). NEFSC CRD 03-03.

Cadrin, S. X. 1999. A precautionary approach to fishery control rules based on Surplus production modeling. NOAA Tech. Memo. NMFS-F/SPO-40. 17-22 pp.

Clark, S.H. 1981. Use of trawl survey data in assessments. *In:* Doubleday, W.G. and D. Rivard (eds.), Bottom Trawl Surveys, p. 82-92, Can. Spec. Publ. Fish. Aquat. Sci. 58.

Conser, R.J. and J.E. Powers. 1990. Extensions 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., 32:461-467.

Delong, A., K. Sosebee, and S. Cadrin. 1997. Evaluation of vessel logbook data for discard and CPUE estimation. NEFSC Laboratory Reference Document 97-xx.

Efron, B. 1982. The jackknife, the bootstrap and other resampling plans. Phila. Soc. for Ind. and Appl. Math. 38: 92 p.

Fournier, D. A., and C. P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39:1195-1207.

Gavaris, S. 1988. An adaptive framework for the estimation of population size. CAFSAC Res. Doc. 88/29, 12 p.

Halliday, R.G. 1973. The flatfish fisheries of the Scotian Shelf. Int. Comm. Northw. Atl. Fish., Res. Doc. 73/103, Ser. No. 3064 (mimeo), 46 p.

Halliday, R.G. 1987. Size and age at sexual maturity of Atlantic argentine, *Argentina silus*: a critique. Environ. Biol. Fishes 19(2):165-261.

Howe, A.B., F.J. Germano, J.L. Buckley, D. Jimenez, and B.T. Estrella. 1981. Fishery resource assessment, coastal Massachusetts. Completion Report, Massachusetts Division of Marine Fisheries, Commercial Fisheries Review Div. Project 3-287-R-3.

Ianelli, J. N., and D. A. Fournier. 1998. Alternative age-structured analyses of NRC simulated stock assessment data. NOAA Tech. Memo. NMFS-F/SPO-30. pp. 81-96.

Lange, A.M.T. and F.E. Lux. 1978. Review of the other flounder stocks (winter flounder, American plaice, witch flounder, and windowpane flounder) off the northeast United States. NMFS, NEFC, Woods Hole Lab. Ref. Doc. No. 78-44, 53 pp.

Mayo, R.K., L. O'Brien, and N. Buxton. 1992. Discard estimates of American plaice, *Hippoglossoides platessoides*, in the Gulf of Maine northern shrimp fishery and the Gulf of Maine-Georges Bank large-mesh otter trawl fishery. Appendix to NEFSC Ref. Doc. 92-07, Res. Doc. SAW 14/3, 40 p.

Methot, R. D. 1990. Synthesis model: an adaptive framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. 50:259-277.

NEFSC [Northeast Fisheries Science Center]. 2002. A Report of the Groundfish Assessment Review Meeting (GARM), October 2002. NEFSC Ref. Doc. 02-16, 511p.

NEFSC [[Northeast Fisheries Science Center]. 2002. Final Report of the Working Group on Re-Evaluation of biological Reference Points for New England Groundfish. NEFSC Ref. Doc. 02-04.

NEFSC [Northeast Fisheries Science Center]. 1994. Report of the 18th Northeast Regional Stock Assessment Workshop (18th SAW), Stock Assessment Review Committee (SARC) consensus summary of assessments. NEFSC Ref. Doc. 94-22, 281 p.

Northern Shrimp Technical Committee. MS 1984. Results of the 1983 northern shrimp survey in the western Gulf of Maine, August 1983. Unpublished Report, Woods Hole, MA, 16 p.

O'Brien, L. and R.W. Brown. 1996. Assessment of the Georges Bank Haddock stock for 1994. NEFSC Ref. Doc. 95-13, 108 p.

Otter Research. 2001. An introduction to AD Model Builder Version 6.0.2 for use in nonlinear modeling and statistics. Otter Research, Ltd. Sidney, B.C. V8L 3S3, Canada.

Parrack, M.L. 1986. A method of analyzing catches and abundance indices from a fishery. Int. Comm. Conserv. Atlantic Tunas, Coll. Vol. Sci. Pap., 24:209-211.

Quinn, T. P., II, and R. B. Deriso.1999. Quantitative fish dynamics. Oxford University Press, New York, 542 pp.

Rivard, D. 1980. APL programs for stock assessment. Can. Tech. Rep. Fish. Aquat. Sci. 953.

SAS. 1985. SAS User's Guide: Statistics, version 5. SAS Institute, Cary, NC. 956 p.

Thompson, W.F. and F.W. 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. Pac. Halibut Comm. 8: 49 p.

Walsh, S.J., R.B. Millar, C.G. Cooper, and W.M. Hickey. 1992. Codend selection in American plaice: diamond versus square mesh. Fish. Res., 13:235-254.

Wigley, S.E. MS 1994. Estimation of foregone yield associated with the discarding of witch flounder *(Glyptocephalus cynoglossus)* in the Gulf of Maine northern shrimp *(Pandalus borealis)* fishery, 1982-1992. M.S. thesis. North Carolina State University, Raleigh, NC. 91 p.

Wigley, S.E. and R.K. Mayo. 1994. Assessment of the Gulf of Maine-Georges Bank witch flounder stock for 1994. Northeast Fisheries Science Center Reference Document 94-17, 86 p.

Wigley, S.E. , J.K.T. Brodziak and S.X. Cadrin. 1999. Assessment of the witch flounder stock in Subareas 5 and 6 for 1999. Northeast Fisheries Science Center Reference Document 99-16, 153 p.