# C. GOOSEFISH

## TERMS OF REFERENCE

The following terms of reference were addressed for goosefish:

1. Summarize results of cooperative NEFSC-industry goosefish survey conducted during 2001.

2. Update fishery-independent information from SARC 31 assessment.

3. Update commercial fishery data, including landings and discard sampling information.

4. Evaluate stock status relative to reference points.

## INTRODUCTION

Goosefish fisheries are managed in the Exclusive Economic Zone (EEZ) through a joint New England Fishery Management Council - Mid-Atlantic Fishery Management Council Monkfish Fishery Management Plan (FMP). The overfishing definition for goosefish is:

> *Monkfish in the northern and southern management areas are defined as being overfished when the three-year moving average autumn survey weight per tow falls below the 33rd percentile of the time series, 1963-1994, or when fishing mortality exceeds* $F_{threshold}$. *Monkfish are in danger of becoming overfished when the three-year moving average autumn survey weight per tow falls below the median of the three-year moving average during 1965-1981 and when fishing mortality is between* $F_{target}$ *and* $F_{threshold}$.
>
> *For the northern and southern areas,* $F_{threshold}$ *is based on conditions of stock stability at high abundance, calculated at the fishing mortality rate that prevailed during 1970-1979.* $F_{target}$ *for the southern area is* $F_{0.1}$. *For the northern area,* $F_{target}$ *is currently undefined.*

There are currently two assessment units for goosefish which are based on differences in the temporal pattern of recruitment (NEFSC survey indices for 10-20 cm goosefish), the spatial and temporal distribution of all sizes of goosefish in NEFSC surveys, perceived differences in growth patterns, and differences in the contribution of fishing gear types (mainly trawl, gill net, and dredge) to the landings. NEFSC surveys continue to indicate different recruitment patterns in the two units in the most recent years. The perceived differences in growth were based on studies about 10 years apart and under different stock conditions (Armstrong (1987): Georges Bank to Mid-Atlantic Bight, 1982-1985; Hartley (1995): Gulf of Maine, 1992-1993). Age, growth, and maturity information recently available from the NEFSC 1992-2001 surveys and the Industry Cooperative 2001 survey now indicate small differences in age, growth, and maturity between the areas. There continue to be significant differences in the contribution of different gear types to the landings. A recent genetics study (Chickarmane et al. 2000) indicated no genetic differences among goosefish collected from North Carolina to Maine in depths up to 300 m.

Because of the uncertainty re. stock structure, this assessment was conducted under the two assessment unit hypothesis and as a combined stock. The preponderance of the biological evidence (recent age, growth, maturity, and genetic information) suggests that use of a single stock hypothesis in the assessment might be appropriate. However, substantial differences in the fisheries exist, and it may be desirable to maintain separate management areas to accommodate these differences.

The research survey strata and statistical areas used to define the northern and southern management regions were as follows:

| Survey | Northern Area | Southern Area |
|---|---|---|
| NEFSC Offshore bottom trawl | 20-30, 34-40 | 1-19, 61-76 |
| ASMFC Shrimp | 1 -12 | |
| Shellfish | 49-54, 65-68, 71-72, 651,661 | 1-48, 55-64, 69-70 73-74, 621, 631 |
| Statistical areas | 511-515, 521-523 561 | 525-526, 562, 537-543, 611-636 |

The southern deepwater extent of the range of goosefish (*Lophius americanus*) overlaps with the northern extent of the range of blackfin goosefish (*Lophius gastrophysus*) (Caruso, 1983). These two species are very similar morphologically, and this may create a problem in identification of survey catches and landings from the southern extent of the range of goosefish. The potential for a problem however is believed to be small. The NEFSC closely examined winter and spring 2000 survey catches for the presence of blackfin goosefish and found none. The cooperative goosefish survey conducted in 2001 caught only 8 blackfin goosefish out of a total of 6,364 goosefish captured in the southern management region.

The spatial distribution of goosefish catches in winter, spring, and autumn bottom trawl surveys and the summer scallop survey is shown in Figure C9. The winter and scallop surveys do not sample in the Gulf of Maine.

Larval distributions have been inferred from collections by the NEFSC Marine Resources Monitoring, Assessment and Prediction (MARMAP) ichthyoplankton survey (Steimle et al. 1999). Larvae were collected during March-April over deeper (< 300 m) offshore waters of the Mid-Atlantic Bight. Later in the year, they were most abundant across the continental shelf at 30 to 90 m. Larvae were most abundant at integrated water column temperatures between 10-16° C, and peak catches were at 11-15° C regardless of month or area. Relatively few larvae were caught in the northern stock area.

## FISHERY DATA

### U.S. Landings

Landings statistics for goosefish are sensitive to conversion from landed weight to live weight, because a substantial fraction of the landings occur as tails only (or other parts). The conversion of landed weight of tails to live weight of goosefish in the NEFSC weigh out database is made by multiplying landed tail weight by a factor of 3.32.

For 1964 through 1989, there are two potential sources of landings information for goosefish; the NEFSC "weight-out" database, which consists of fish dealer reports of landings, and the "general canvass" database, which contains landings data collected by NMFS port agents (for ports not included in the weight-out system) or reported by states not included in the weight-out system (Table

C1). All landings of goosefish are reported in the general canvass data as "unclassified tails." Consequently, some landed weight attributable to livers or whole fish in the canvass data may be inappropriately converted to live weight. This is not an issue for years 1964 through 1981 when only tails were recorded in both databases. However, for years 1982 through 1989, the weight-out database contains market category information which allows for improved conversions from landed to live weight. The two data sources produce the same trends in landings, with general canvass landings slightly greater than the weight-out system. It is not known which of the two measures more accurately reflects landings, but the additional data sources argue for use of the general canvass landings for years 1964 through 1981 while market category details available in the weight-out system argue for use of this database for years 1982 through 1989. Until the mid-1970's, many of the goosefish caught were sold outside of dealers or used for personal consumption, introducing further uncertainty into the early estimates of landings.

Beginning in 1990, most of the extra sources of landings in the general canvass database were incorporated into the NEFSC weight-out database. However, North Carolina reported landings of goosefish to the Southeast Fisheries Science Center and until 1997 these landings were not added to the NEFSC general canvass database. Since these landings most likely come from the southern management region, they have been added to the weight-out data for the southern management region for 1977-1997 (Table C1).

Beginning in July 1994, the NEFSC commercial landings data collection system was redesigned to consist of vessel trip reports (VTR data) and dealer weigh-out records. The VTRs include area fished for each trip which is used to apportion dealer-reported landings to statistical areas. Each VTR trip should have a direct match in the dealer data base; however, this is not always true. For data with no matches, we dropped the record if there was a VTR with no dealer landings and retained the record if there were dealer landings but no VTR. For dealer landings with no matching VTR, we apportioned the landings to area using proportions calculated from successfully matched trips pooled over gear, state and quarter.

Total landings (live weight) remained at low levels until the middle 1970s, increasing from hundreds of metric tons to around 6000 mt in 1978 (Table C1, Figure C1). Landings remained stable at between 8,000-10,000 mt until the late 1980s. Landings increased steadily from the late 1980s through 1992, and have fluctuated around 26,000 mt since 1993. Peak landings occurred in 1997 (28,517 mt) and have declined slightly since then. By region, landings began to increase in the north in the mid-1970s, and began to increase in the south in the late 1970s. Most of the increase in landings in recent years has been from the southern region.

Trawls, scallop dredges and gill nets are the primary gear types that land goosefish (Table C2, Figure C2). During 1998-2000, trawls accounted for 54% of the total landings, scallop dredges about 17%, and gill nets 29%. In recent years trawl landings (mt) are greater in the northern than southern areas, while scallop dredges and gill nets have landed more from the south than from the north.

Until the late 1990s, total landings were dominated by landings of goosefish tails.

From 1964 to 1980 landings of tails rose from 19 mt to 2,302 mt, and to 7,191 mt in 1997 (Table C3). Landings of tails have declined since 1997 (to 3,582 mt in 2000), while landings of gutted whole fish have increased steadily. On a regional basis, most tails were landed from the northern component in the 1960's (75 to 90%) through to the late 1970's (74% in 1978) (Tables C4, C5). From 1979 to 1989, landings of tails were about equal from both regions. In the 1990's, landings of tails from the south began to predominate, providing 60% or more of tails. In 2000, landings of tails from the two areas were approximately equal.

Beginning in 1982, several market categories were added to the system (Table C3). Tails were broken down into large (> 2.0 lbs), small (0.5 to 2.0 lbs), and unclassified categories and the liver market category was added. In 1989, unclassified round fish were added; and in 1991, peewee tails (<0.5 lbs) and cheeks appeared. Finally, in 1992 belly flaps were also recorded. Whole gutted fish were first recorded in 1993.

Goosefish livers have become a very valuable product. Landings of livers increased from 10 mt in 1982 to an average of over 600 mt during 1998 - 2000. During 1982-1994, ex-vessel prices for livers rose from an average of $0.97/lb to over $5.00/lb, with seasonal variations as high as $19.00/lb. Landings of unclassified round (whole) or gutted whole fish jumped in 1994 to 2,045 mt and 1,454 mt, respectively; landings of gutted fish continued to increase through 2000. The tonnage of peewee tails landed increased through 1995 to 364 mt and then declined to 153 mt in 1999 and 4 mt in 2000 when the category was essentially eliminated by regulations.

Foreign Landings

Landings (live wt) from NAFO areas 5 and 6 by countries other than the US are shown in Table C1 and Figure C1. Reported landings were high but variable in the 1960s and 1970s with a peak in 1973 of 6,818 mt. Landings were low but variable in the 1980s, declined in the early 1990s, and have been below 200 mt in recent years.

Size Composition of U.S. Landings and Catch

Table C6 shows the number of commercial samples taken through the port sampling program for 1996-2000. Length frequencies of the samples are shown in Figure C4; these were expanded to landings using the length-weight equations in Almeida et. al. (1995) (Figure C5). In 1996 “unclassified round” landings from the south were expanded using the “unclassified round” samples (n=2) from the north. In 1997 there were no samples for “tail only”, so landings in this market category were distributed according to the proportion of peewee, small and large tail landings within each stock area. Sampling intensity and coverage was low in 1998. Length frequency of landings for unsampled market categories was estimated according to the proportion of peewee, small, and large tail landings in the north and large and small tails in the south. In 1999 “tail small” was used to expand “tail peewee” landings within each stock component. “Head on gutted” was used for unclassified round, and “tail only” landings were redistributed according to the proportion of small and large tail landings. In 2000, sampling increased but sampling intensity varied widely among market categories and ports.

Length composition data collected by the NEFSC fishery observer program (sea sampling data) were summarized for 1996-2000. Sea sampling data for goosefish were collected aboard trawls, scallop dredges and gill nets (drift and sink). Figures C6 and C7 show length frequency distributions from sea sampling data by major gear type, stock region and year. Discards were generally between 20-40 cm, while kept fish were greater than 40 cm.

Discard Estimates

Catch data from the fishery observer and VTR databases were used to investigate discarding frequencies and rates. The number of tows or trips with goosefish discards available for analysis varied widely among stocks and gear types (Tables C7 and C8). Discard ratios (kg discarded / kg kept) from the two data sources were consistent (Figure C8). Scallop dredges generally had the highest discard ratio while gill nets had the lowest. The most frequent reasons for discarding in the trawl and scallop fisheries were that the fish were too small, either for the market or for regulations. In the gill net fisheries, poor quality was the primary reason for discarding.

We estimated annual mt of goosefish discarded by calculating discard ratios from the observer program on a management region, gear type and half-year basis. We applied the discard ratios to reported landings (live weight, by stock, gear type and half-year cells) to derive metric tons discarded and total catch (Tables C9 and C10). If no sampling data were available for a cell, we applied the overall mean discard ratio for all gears and years. The overall annual discard ratio (Table C10) ranged from 0.07 - 0.27 mt discarded per mt kept. The percentage of the catch discarded has ranged from 6-21%, with the highest rates occurring in 2000.

Catch per Unit Effort by Gear and Depth

Commercial catch per unit effort (CPUE) from the VTR database was examined by gear type in order to determine if a depth effect was present, especially in the deepest waters. Scallop dredge, large and small mesh gill net, and otter trawls were examined separately. Depth zones were categorized in 20 fathom increments starting with 0-20 fathoms (zone 1) and ending with zone 10 (greater than 180 fathoms). Obvious outliers were removed before analysis based on examination of the actual logbooks.

Table C11 presents the number of observations, median CPUE by depth zone and the estimated depth effect from a generalized linear model incorporating year, quarter, vessel ton class and depth zone. Dredge gear does not fish in deep waters and does not show changes in CPUE with depth. Large and small mesh gill nets fish in deeper waters, but do not show a trend in CPUE with depth. In contrast, trawls fish in deep waters and show an increasing trend in CPUE with depth. However, this apparent trend is due to a loss of low CPUE values at greater depths; maximum catch rate is consistent over all depths. Examining only directed trips (trips in which at least half of the catch (kg) was goosefish) removes the apparent trend with depth by removing most of the low catch rates in shallow water (Table C12). Thus catch per unit effort does not appear to have a depth effect associated with any gear. However, the low sample sizes in the deepest water do not allow definite conclusions to be reached.

During the examination of catch rates by depth, it was observed that few trawl trips fall into the directed category, as defined above. Table C13 shows the number of directed and total trips by gear and stock area and the associated landings. Although trawl trips are infrequently directed in both the north and south, 6.1% and 8.8%, respectively, the proportion of catch associated with these trips is much higher in the south, 24% north and

76% south. This difference between north and south was not apparent in either gill net fishery.

Selectivity of Trawls and Scallop Dredges

An exploratory analysis of selectivity patterns of trawls and scallop dredges was performed for SARC 31 (NEFSC 2000). The analysis was based on the following assumptions:

1) The index of abundance in a given length category is proportional to the population. That is, $n_i = c\ N_i$, where $c$ is a constant of proportionality over all length categories and years, and $n_i$ and $N_i$, respectively, are the abundance index and population size of the $i$th length category.

(2) The proportion of the population vulnerable to the fishing gear (vulnerability) is an S-shaped function of length, which can be described by a half-gaussian curve:

$$v_i = \exp[-0.5(l_i - L_{full})^2/s], \text{ if } l_i < L_{full}$$
$$= 1, \qquad \text{if } l_i \geq L_{full}$$

where $l_i$ is the length of the ith category and $L_{full}$ is the length of fully vulnerable individuals.

(3) The exploitation rate ($u$) operates equally on all vulnerable individuals in the population, and thus, the catch in number of the ith length category is

$$C_i = u\ v_i\ N_i.$$

The length-frequency distributions in proportion ($p_i$) are then expressed by the equations in assumptions (1) and (3):

$$p_i = C_i / \Sigma\ C_i = v_i\ n_i / \Sigma\ v_i\ n_i.$$

If $P_i$ is the observed proportion of catch in the ith length category, which is a measurement of population's $p_i$ with an error of $e_i$, it implies that $P_i = p_i + e_i$.

The method of least squares was used to estimate the location parameter $L_{full}$ and the shape parameter $s$ of the vulnerability, or selection, curve. In order to apply the method, the number of samples for the abundance index should be sufficient, i.e. the values of $n_i$'s of all length categories should be large enough to make a smoothed length-frequency distribution without too many null categories. Gillnets were not included in the analysis because the upper range of survey length-frequency distributions does not extend to that sampled from the gillnets.

For the northern stock, the vulnerability of kept goosefish sampled from vessels using scallop dredges was consistent during 1996-1998, with less than 10% vulnerable at 40 cm and almost 100% vulnerable near 45 cm. Vulnerability curves of kept goosefish from trawlers were similar in 1997 and 1998 but different from that in 1996 (Table C14). Some discards in 1996 may have been mis-coded as kept, resulting in a less steep curve.

For the southern stock, the vulnerability of kept goosefish to trawls and scallop dredges was similar in 1996 and 1997, when compared with data from scallop and winter surveys (Table C14). Differences occurred after 1998 although some were similar. It should be noted that relatively small samples were collected in 1998-1999 compared to 1996-1997. The small samples probably biased the length-frequency distributions of the kept portion of the catch.

## RESEARCH SURVEY ABUNDANCE AND BIOMASS INDICES

NEFSC Survey Indices

NEFSC spring and autumn bottom trawl survey indices were standardized to adjust for statistically significant effects of trawl type and vessel on catch rates as noted below. The trawl conversion coefficients apply only to the spring survey during 1973-1981.

| Effect | Coefficient | Source |
|---|---|---|
| Trawl | Weight: 0.2985<br>Number: 0.4082 | Sissenwine and Bowman, 1977 |
| Vessel | Weight: Not significant<br>Number: 0.83 | NEFSC, 1991 |

Northern Region

Indices from NEFSC autumn research trawl surveys indicate that biomass fluctuated without trend between 1963-1975, appears to have increased briefly in the late 1970's, but declined thereafter to near historic lows during the 1990's. In 2000 the index increased to its highest level since 1984 (Table C15, Figure C10). The three year moving average of the index (1998-2000) is currently at 57% of the 1965-1981 biomass target (Table C49). Abundance in numbers (Table C15, Figure C11) declined during the early 1960s, and then fluctuated without trend until the late 1980s. Abundance increased steadily from the late 1980s to a peak in 1994, declined to 1997, increased in 1998 and 1999 and increased sharply in 2000. The 2000 point estimate for numbers is the highest in the series.

Indices from the NEFSC spring research trawl surveys reflect similar trends of relatively high biomass levels in the mid 1970s (but with possible declines in the late 1970s), a declining trend from the early 1980s to the lowest values in the time series in 1998 and an increasing trend since then (Table C16, Figure C12). As in the autumn survey series, abundance in numbers fluctuated until the early 1980s (Table C16, Figure C13). Since 1996, numbers have trended upwards and reached the highest levels in the time series in 2000 and 2001. Figure C14 shows the fall and spring survey indices plotted together for comparison of trends.

Other surveys conducted in the northern management region cover shorter periods of time and/or smaller portions of the region, and are not included in this assessment because of their limited coverage. For example, the NEFSC sea scallop survey in the northern goosefish management region includes only a few strata on the northern edge of Georges Bank and the ASMFC shrimp survey covers only the western Gulf of Maine.

Length distributions have become increasingly truncated over time (Figure C15). By 1990, fish greater than 80 cm long were uncommon in length frequency distributions, and by 1996, fish greater than 60 cm had become relatively uncommon as well. The minimum, mean and maximum lengths in the trawl surveys have declined steadily over time (Figure C16).

Several modes potentially representing strong year classes have appeared consistently in survey distributions in recent years. Abundance indices for goosefish 10-20 cm TL (corresponding approximately to age 1 goosefish) were estimated to help identify potential recruitment patterns (Figure C17,

Table C17). To the extent that these indices reflect recruitment, recruitment in the northern area has increased in the past decade. Relatively strong year-classes were produced in 1992, 1993, 1998 and 1999. Length frequencies and survey abundance at age data corroborate the suggestion of relatively strong 1998 and 1999 year-classes (Figure C15, Table C18) in the northern area.

Survey age data are available for 1993-2000 from the autumn trawl survey and for 1995-2001 for the spring trawl survey. The mean length at age is shown in Table C18 and Figures C18-C20). Within the range of ages observed in the surveys, growth is essentially linear and there are no obvious differences with gender or stock. The stratified mean number per tow at age is shown in Table C19.

Southern Region

Biomass indices from the NEFSC autumn research survey declined rapidly in the second half of the 1960s, and then fluctuated until the early 1980s (Table C20, Figure C21). In the mid-1980s, biomass declined and has remained low since 1987. The three year moving average of the index (1998-2000) is currently at 23% of the 1965-1981 biomass target (Table C49). Abundance in numbers shows similar declines after the mid-1960s, with a spike in 1972, slight increases in the late 1970s-early 1980s and a decline thereafter (Figure C22). In recent years, abundance in numbers has fluctuated without trend at low levels.

The Overfishing Definition biomass target and thresholds for the southern component are based on NEFSC autumn survey indices beginning in 1963. NEFSC survey strata south of Hudson Canyon were not sampled during 1963-1966, and so indices for those years are not directly comparable to indices for 1967 and later years. SARC 31 recommended the adoption of southern component biomass target and thresholds based on indices for 1967-1981 and 1967-1994, respectively. This revision changes the biomass target from 1.848 kg per tow to 1.846 kg per tow, and the biomass threshold from 0.750 kg per tow to 0.704 kg per tow.

The NEFSC spring research survey data reflects similar trends as the autumn series: stock levels remained fairly high during the mid 1970s - early 1980s, but declined to record low levels in the early 1980s and have fluctuated at low levels in recent years (Table C21, Figures C23 and C24).

Indices based on the NEFSC winter flatfish survey have fluctuated without trend, consistent with lack of trend in other surveys during 1992-2001 (Table C22, Figures C21, C23, C29); however, the 2001 biomass index was the highest in this series. The abundance index did not increase to a similar degree. Age data are available for the winter survey for 1997-2001 (Table C23, Figure C27). The mean length at age for the winter survey samples is similar to mean length at age from NEFSC spring surveys (Figure C20).

Abundance indices based on the NEFSC sea scallop survey show an increasing trend during 1984-1994 followed by a rapid decline from 1994-1998; however, the abundance index increased in 1999 (Table C24, Figure C28). Finfish data for scallop surveys conducted during 2000 and 2001 are not yet available.

Figure C29 compares biomass and abundance indices from all NEFSC surveys in the southern management region.

Length distributions from the southern region show increasing truncation over time (Figure C30), which is reflected in declines in minimum, mean and maximum length over time (Figures C31 and C32). Maximum lengths declined by approximately 20 cm or more over the time series.

As in the northern region, recent year class events are rarely observable in survey length frequency distributions at lengths greater than 40 cm, Currently, fish greater than 60 cm are rare, especially when compared to the 1960s. Any recent strong recruitment does not appear to survive long enough to contribute substantially to increased stock biomass.

Management Areas Combined

Tables C25-C27 and Figures C33-41 present survey information from the fall, spring and scallop surveys for the northern and southern management regions combined.

MA DMF Survey Indices

Surveys conducted by the Massachusetts Division of Marine Fisheries show trends in biomass and abundance broadly similar to NEFSC surveys in the northern region (Figure C42). Biomass indices for the state waters north of Cape Cod show a declining trend in both the spring and the fall. Abundance indices fluctuated at low levels until the 1990s when there was a small peak in 1991 and a large spike in 1995. Abundance of goosefish in inshore waters appears lower during the spring; however, the highest point in the spring series is also 1995. A peak in abundance was observed in 1994 in the NEFSC fall survey. The MADMF index shows an increase in biomass in 2000, but does not indicate the increased abundance in 2000 that the NEFSC survey index does.

In Massachusetts waters south of Cape Cod, biomass indices have remained at or near their lowest levels since around 1990 and abundance has been consistently very low.

## 2001 COOPERATIVE GOOSEFISH SURVEY

Methods

A directed survey for goosefish was conducted in cooperation with the fishing industry during Feb 27 -May 17, 2001. The F/V Drake (87 ft. trawler, home port Portland, ME) and the F/V Mary K (96 ft. trawler, home port New Bedford, MA) were chartered to conduct the survey. The Drake had two nets which were alternated depending on bottom type (Figure C43); the Mary K used one net for all tows (Table C28). The Drake sampled the Gulf of Maine and Georges Bank, the Mary K sampled southern New England and the mid-Atlantic shelf down to Cape Hatteras.

The basis for the survey was a stratified random design with sampling effort proportional to reported fishing effort during 1995-1999. Additional station locations were assigned by fishermen. The stratum boundaries were those used in NEFSC bottom trawl surveys (defined by depth), with an additional set of strata from Georges Bank south in 100 to 500 fathoms. The realized distribution of the 284 survey stations successfully occupied is shown in Figure C44. The survey stations were completed during Feb. 27 to April 6.

Standard operating procedures were followed by each vessel. These specified such variables as tow time, tow speed, scope ratios, sampling protocols, etc. Ancillary data collected for each tow included bottom contact time, measured using an inclinometer hung from the footrope of the net, boat position (GPS), and temperature. The electronic data were collected at intervals ranging from 1 to 6 seconds; clocks were synchronized among the sensors. Survey catches were processed using standard procedures for NEFSC surveys.

In addition to the survey stations, 64 tows were conducted for mensuration of the three nets, efficiency estimation, inter-net and inter-vessel calibration, video work and to examine the outer depth distribution of goosefish (Table C29). The tows were conducted in waters off southern New England (Figures C45-C50).

The net mensuration work was done using a NetMind trawl mensuration system for measuring wing spread, door spread, and headrope height on both vessels. The general protocol was to conduct a pair of 30-minute tows at approximately 40-fathom depth increments (30-150 fathoms for the Drake and 30-280 fathoms for the Mary K, Figure C45). The second tow of each pair was fished in the opposite direction of the initial tow to account for variation in tow direction relative to current direction. Nets were set and towed along the depth contour.

To compare catch rates between net types on the Drake, a series of tows done on soft bottom with net 1 were repeated using net 2 (Figure C46). The tows with net 1 were completed on May 11 and the tows with net 2 on May 12-13. Tows were done at 40, 70, 100 and 140 fathoms. Repeated tows were adjacent to the first tow, not on the original tow path.

A series of depletion tows were conducted on the Mary K and the Drake (net 1) to estimate absolute efficiency of the gear (Figure C47). Standard 30-minute tows were repeated in alternate directions along a single tow path (different tow paths for each vessel) until the catch rates dropped to zero or near zero.

Comparisons between the two vessels (Drake using net 1) were made by conducting a series of paired tows in which the vessels fished next to each other at tow locations in depths ranging from 30 to 140 fathoms (Figure C48). In another set of experiments, the Mary K repeated 7 tows completed by the Drake about 5 days earlier (Figure C49). These experiments were not used in estimating biomass and population size, but provided a direct estimate of relative performance of the two vessels and nets.

Video camera observations were made using an underwater camera system to evaluate the catchability of goosefish by the three nets used in the cooperative survey (Figure C50). The video tapes were used to examine the behavior of goosefish as they encounter the gear, to assess the degree to which herding occurs and to obtain a qualitative sense of the efficiency of the gear. A third wire camera system was mounted on the headrope of the net and videos were viewed in real time and recorded. The camera system's pan and tilt unit allowed the operator to change the field of view of the camera and thus view separate areas of the net (i.e. wings, center of sweep, groundcables) to provide a broader understanding of goosefish behavior in response to the gear. Camera tows were conducted in daylight in water depths of 27-37 fathoms. The net was towed with the codend open until the scientists and fishermen felt they had enough video data to adequately describe the behavior of the goosefish within the trawl.

Area swept biomass and population numbers were estimated for each survey tow. The distance covered by each tow was estimated from bottom contact time (based on inclinometer data) and speed of the vessel as derived from GPS position data during bottom contact (Figures C51-C54). Width of the tow path for each tow was estimated from wingspread-depth relationships developed from the mensuration work (Figures C55-C57). Where inclinometer data were missing for a tow, we adjusted nominal tow distances according to inclinometer:nominal distance relationships from tows with high quality sensor data. For the Mary K, this relationship was depth-dependent (Figure C58). Where GPS data were missing, we used average speed from tows with good quality sensor data (by vessel) to calculate the distance covered. A second set of area swept estimates was derived using nominal tow distance (distance covered in the time between winch lock and re-engage) for the Mary K because it is uncertain how much of the bottom contact time after winch lock is actually fishing time (with the net moving forward).

To estimate population biomass (numbers), we calculated goosefish densities in each stratum as the sum of the numbers caught divided by the sum of the area swept. Biomass in each stratum was estimated as the product of number of fish and mean weight of fish in the stratum. Biomass and numbers were summed over strata to arrive at minimum biomass and population size. Biomass and population size were also estimated under a range of assumptions regarding net efficiencies. The efficiency assumptions were derived from the depletion and calibration experiments. We used the depletion experiments to estimate efficiency of the Mary K's net and the Drake's net 1. The Drake's net 2 was adjusted to the Drake's net 1 based on the paired tow experiments.

## RESULTS - COOPERATIVE GOOSEFISH SURVEY EXPERIMENTAL TOWS

Results of the Drake net calibration experiments are summarized in Table C30 and Figure C59. There was not a strong correspondence between catch rates with the two nets, but net 2 tended to catch slightly less than net 1. We used the overall ratio of net 2 : net 1 catches (0.92) as the estimate of efficiency of net 2 relative to net 1.

The paired tows between the Drake (net 1) and the Mary K were analyzed under both assumptions regarding tow distance for the Mary K (inclinometer distance estimates, nominal distance estimates) (Figures C60 and C61, Table C31). Assuming inclinometer distances for both vessels, the ratio of numbers per $nm^2$ Drake:Mary K was 1.10; assuming nominal distances for the Mary K brought the ratio to 0.93. The repeated tow experiments indicated Drake:Mary K ratios for numbers per $nm^2$ of 0.76 - 0.88 (Figure C62).

The video footage provided no evidence of herding of goosefish by the gear, nor of strong escape responses. Goosefish generally were not visible before being contacted by the tickler chain, but when hit by the chain would flip up into the water column and then drift passively into the net.

## RESULTS - COOPERATIVE GOOSEFISH SURVEY

A total of 310 survey tows were completed during the project. Of these, 284 tows had no gear problems or other major difficulties, and could be used to estimate goosefish abundance (125 tows in the northern management region, 159 tows in the southern management region). Over 9,000 goosefish (16,500 kg) were caught during the survey. More than 3,000 of the goosefish were sampled for age and sex determination, maturity, and food habits. The size of goosefish ranged from 13 cm to 110 cm; ages ranged from 2 to 10 years.

Eight blackfin goosefish were caught in the southern management region (Figure C63). Their identification was later confirmed by systematists at Harvard's Museum of Comparative Zoology (K. Hartel, personal communication).

Nine incidences of cannibalism by goosefish were recorded (Table C32). The evidence of cannibalism ranged from goosefish skeletal remains in stomachs to partially digested goosefish. One stomach contained two goosefish prey. Size of the cannibals ranged from 63-105 cm, all were female; sizes of the prey were 45-49 cm.

Length-weight relationships for male and female goosefish by management area and the entire region are shown in Figure C64. Females in the south appear to be heavier for a given length after reaching about 60 cm total length; however this is likely due to the advanced stage of gonadal development in many of the females sampled in the southern region. In 96 females from the southern region whose gonads were weighed, an average of 27% of the total body weight was egg veil.

Mean length at age by sex and management area are shown in Figure C65. Differences in growth between males and females are undetectable before age 7, when growth in males appears to slow, while female growth continues to increase almost linearly. Few males greater than 65 cm (predicted age 7) were captured. Mean length at age by region and for sexes combined is shown in Figure C66 and Table C33. Size at age was slightly higher in the southern management region. This is consistent with seasonal changes in growth seen in NEFSC survey data for goosefish. Mean weight at age (Figure C67) increases exponentially up through the oldest ages observed in the survey (10 years).

Sex ratios at length (Figure C68) indicate that in both management regions, all individuals larger than about 70 cm are female. In the north, sex ratios average around 50:50 for goosefish 20-60 cm. In the south, sex ratios are about 50:50 for goosefish 20-40 cm total length; for goosefish 40-60 cm, the percent of females drops to 30-40%, and thereafter rises to 100% females by around 70 cm.

Maturity ogives for females and males were fit using probit analysis (Figures C69 and C70). Fifty percent of females are mature at 40 cm (4.7 years) in the northern region and at 46 cm (5.1 years) in the southern region. The estimates of 50% female maturity for regions combined is 43 cm and 4.8 years. These estimates correspond closely with other studies conducted using macroscopic inspection of female gonads; however a study done using histological methods indicated a higher size at 50% female maturity (57 cm, Martinez 1999). Fifty percent maturity for males is estimated to be 35 cm (4.1 years) in the northern region and 37 cm (4.3 years) in the southern region (Figure C70). For regions combined, 50% of males are estimated to be mature at 36 cm (4.2 years).

Swept area biomass and population size estimates under varying assumptions about net efficiencies (Table C34) and tow distance for the Mary K are shown in Table C35 and

Figure C71. Minimum estimates (assuming 100% efficiency of nets) range 64-72 thousand metric tons and 43-48 million goosefish for both areas combined. The range in these estimates is due to the method of estimating distance towed by the Mary K (nominal vs. inclinometer distances).

The length composition of the monkfish population estimated from the cooperative survey (based on minimum population size and assuming inclinometer distances for all nets) is shown in Figure C72. In both management regions, most of the population is less than the minimum landing size required under the FMP. Length frequencies from the NEFSC winter survey for 2001 are very similar to the length frequencies derived from the cooperative survey (Figure C73). Minimum spawning biomass was estimated under the inclinometer distance assumption from numbers at length in each management region, sex ratio at length, maturity at length and the length-weight relationship from the cooperative survey samples (Figure C74).

Age composition of the goosefish population by management region and areas combined (Figure C75) was derived from the age-length key for areas combined applied to the number of goosefish at length (by region and for areas combined).

## RELATIVE PRECISION OF F/V COOPERATIVE SURVEY AND COMPARISONS WITH NMFS RESEARCH TRAWL SURVEY

The precision of abundance estimates is an important aspect of research surveys. When the underlying assumptions of stratified random surveys are satisfied, such surveys can provide valid inferences about the true population densities. This section provides estimates of the relative precision of stratified random surveys using the sampling theory summarized in Cochran (1977). The applicability of standard sampling theory to fish populations has been the subject of considerable debate, particularly with respect to the alternatives of model-based estimates (e.g., Pennington 1983, 1986) or explicit spatial models (e.g., Conan and Wade 1989). The choice of design vs model-based methods of estimation usually is motivated by the presence of high variation in the observed catch data . Conventional estimates of the precision, e.g., the standard error of the estimate, can lead to confidence intervals with negative lower bounds. Model-based estimators of abundance account for such variations by assuming a particular statistical model (e.g. lognormal, poisson or delta distribution) for the underlying distribution of the resource. Subsequent inferences are therefore conditional on the validity of the assumed model. Smith (1990) and Myers and Pepin (1990) demonstrated that model-based estimates can result in biased estimates of population means and variances when the underlying model is not supported by the data.

Alternatively, bootstrap resampling methods may be used to estimate the relative precision in complex survey designs (Smith 1997). The bootstrap approach avoids the need to explicitly choose (and justify) an underlying statistical distribution, and it leads to a realistic characterization of the sampling distribution of the mean and variance estimates. This section relies heavily on the theory and applications described in Cochran (1977), Smith (1996, 1997, 2000) and Smith and Gavaris (1993). All of the computations of design efficiency and bootstrap estimators were conducted in Splus using a library of functions written by Stephen Smith, DFO, Dartmouth, NS.

## Methods

Estimates of the mean, standard error, and effective degrees of freedom for stratified random surveys were based on standard equations in Cochan (1977). Under the assumption that the stratified mean would exhibit a Student's t distribution under repeated sampling, an approximate parametric confidence interval for the mean can be constructed. The relative efficiency of the design can be computed by comparing the variance of the stratified random estimate with that which would be obtained under simple random sampling. The computation of a simple random sampling variance for data collected in a stratified random survey is easily computed but complicated (see Smith 2000, Eq. 6).

As shown in Smith and Gavaris (1993), the reduction in variance associated with the use of a stratified random design can be decomposed into two components related to the allocation of samples to strata, and the differences among stratum means. The contribution associated with differences in stratum means is always positive. In contrast, inappropriate allocation of samples to strata can lead to negative values, such that the variance of a stratified random design can be greater than a simple random sample. Such differences can occur when the overall design targets another species or when the survey design reflects a compromise among many target species. Finally, it is possible to estimate a minimum variance that would be obtainable under optimal allocation. Optimal allocation of samples is based on the relative size of the strata and the estimated stratum variance. Minimum variance estimates are useful when contemplating revisions to future sampling designs, and as metric of evaluating the relative efficiency of the realized survey.

Bootstrap sampling of complex survey designs is complicated by the known bias properties of stratified variances of the mean (Rao and Wu 1988). Smith(1997) applied the so-called "mirror-match" of Sitter (1992) to reduce the bias associated with bootstrap variance estimates from small samples. Essentially this approach randomly uses $n_h$ and $n_h$ -1 resamples from each stratum when deriving the bootstrap values. Confidence limits are derived using percentile methods. Smith (1997) demonstrated that this method of computing was preferable to other methods that attempt to correct for differences between the point estimate and the median of the sampling distribution.

## Comparisons with NMFS R/V Trawl Surveys

The results of the cooperative survey were compared to spring, autumn and winter NMFS trawl surveys. Comparisons were made with the most recent NMFS survey and with an additional year, selected for its low mean catch of goosefish. The surveys compared were spring (2001, 1987), fall (2000, 1997) and winter (2001, 1998). Fall and spring surveys were analyzed for the northern and southern management regions and for regions combined. The winter survey does not sample the northern strata, so was analyzed for the southern region only. Catch estimates of monkfish in the NMFS surveys were adjusted to a standard area swept, defined by net width, standard tow duration and standard towing speed. Individual variation in tow distance could not be adjusted for because detailed data on gear performance (e.g. actual bottom contact time) is not available for all surveys.

The cooperative survey results were also analyzed for northern, southern and combined regions. The southern strata in the cooperative survey were reduced to coincide exactly with NMFS survey coverage. This eliminated the deepwater strata and was done to provide comparable strata sets for the comparisons. The response variable for each survey was either the number or weight (kg) caught per tow. To account for differences among the three nets used in the cooperative survey, the catch rates were adjusted according to a range of assumptions re. variations in net width with depth, distance covered during net deployment, and adjustments for estimated contact time (Tables C36 and C37).

Results

Mean catch rates per tow in the cooperative survey were much greater than those observed in the NMFS surveys (Tables C36 and C37). These differences reflect smaller net width and lower efficiency of the rollers on the NMFS fall and spring surveys. The coefficient of variation (CV) of catch rates in the cooperative surveys ranged from 4 to 7%, suggesting a high degree of precision. The NMFS winter survey had CVs about twice as large (11-14%). CV's for the NMFS fall and spring surveys varied from 15% to 50%.

The cooperative survey achieved variance reductions ranging from 50 to 86% over simple random sampling. Most of the gain in precision was attained through stratification, rather than allocation. This suggests that the survey strata were appropriate for the cooperative survey, and that the variations in sample allocations to strata were less important. While the survey strata were also appropriate for the NMFS survey, the allocation of samples to these strata often resulted in reduced precision. In 8 of the 12 comparisons for the spring and autumn surveys, the negative effect of sample allocation resulted in higher variance than would have been obtained via simple random sampling. This inefficiency in allocation for goosefish probably results from an overall allocation scheme for NMFS surveys which targets a wide range of species.

Bootstrap estimation of confidence limits (Tables C38 and C39, Figures C76-C78) resulted in a slight reduction in the length of the interval and provided a non-parametric estimate of the sampling distribution percentiles. No strong evidence of bias (i.e., difference between the bootstrap estimate and the point estimate) was evident for either the mean or variance. Side-by-side comparisons of the parametric and bootstrap confidence intervals revealed only slight differences (Table C40). The length of the putative confidence interval (upper-lower estimates) was slightly larger for the cooperative survey bootstrap estimates and slightly smaller for the NMFS survey bootstrap estimates. The near equivalency of the bootstrap and design-based estimates contrasts with other applications of Smith's methodology (eg. Smith 1996, 1997), and is perhaps due to the spatial dispersion of monkfish. None of the surveys observed wide variations in the number of monkfish per tow as compared with other groundfish and pelagic species. This may reflect a relatively uniform spacing of monkfish in areas of suitable habitat.

Comparisons between the cooperative survey and NMFS winter survey results are highlighted in Table C41 and Figure C79. For this comparison, the cooperative survey was restricted to the strata sampled by the NMFS winter survey in 2001. The estimates from the NMFS winter survey are less precise than the cooperative survey's, but are still considered very good for a multispecies resource survey. Revision of sample allocation in the winter survey could improve the survey's precision for goosefish. However the biggest contrast in the survey estimates is the difference in the total biomass estimates. The ratio of these estimates, assuming that variations in net width and tow path duration have been properly accounted for, suggests that the NMFS winter trawl is about half as efficient (i.e., probability of capture given encounter) as the "average" commercial net. As an exploratory calculation, the distribution of bootstrap estimates of biomass for the winter survey were rescaled to the mean of the cooperative survey. The results, shown in the lower panel of Figure C79 illustrate the that the winter survey has precision comparable to that observed in the cooperative survey.

The net used in NMFS fall and spring surveys appears to be less efficient compared to the cooperative survey nets but more detailed examination is necessary. In particular, analysis of differences between catch rates with the large roller net used by the F/V Drake in the Gulf of Maine and the NMFS survey nets would be instructive.

## EGG PRODUCTION INDICES FROM NEFSC SURVEY LENGTH COMPOSITION DATA

NEFSC survey indices were used to develop indices of egg production. Composite length frequencies, based on a five year summation of catch per tow at length, Ī(L,t) were multiplied by predicted eggs at length Egg(L) and the fraction mature (PMAT(L)). The computational formula is:

$$SSB(t) = \sum_{L} SSB(L,t) = \sum PMAT\ (L) * Eggs(L) * \bar{I}(L,t)$$

where

$$PMAT\ (L)\ =\ \frac{1}{1 + e^{13.9568 - 0.03862325L}}$$

Parameters for PMAT(L) were derived by fitting the logistic function to derived percentiles of fraction mature described in Hartley (1995). The fecundity-length relationship was obtained from Armstrong (1987).

$$Eggs\ (L) = 0.0683\ L^{3.74}$$

$$L = length(mm)$$

Results for the indices of egg production (Figures C80-C82, Table C42) mirror the progressive decline in mean length and have declined steadily over the past two decades.

Currently, about 13% of SSB is produced by fish less than $L_{50}$. In the north, about 10-13% of the egg production is by the partially mature component of the length distribution; in the south, 13-17% of the spawning stock biomass is from the partially mature component of the length distribution.

## ESTIMATION OF MORTALITY AND STOCK SIZE

Natural Mortality Rate

The instantaneous natural mortality rate for monkfish is assumed to be 0.2, based on an expected maximum age of 15-20 years given previous studies of age and growth (Armstrong 1987, Armstrong et al. 1992, Hartley 1995).

Mortality estimates from NEFSC Surveys

Instantaneous total mortality rates (Z) for goosefish were estimated using a length-based method by Beverton and Holt (1956):

$$z = \frac{K(L_\infty - \bar{L})}{(\bar{L} - L')}$$

where K and $L_\infty$ are from von Bertalanffy growth models and LBar is the mean length of individuals in the region (as stratified delta mean catch per tow at length, adjusted for trawl and vessel effects, when significant). $L'$ is the smallest fully recruited length, and was estimated from inspection of LOWESS smoothed length frequency data (Cleveland, 1979) . The value of L' established in the SAW 31 assessment was 30 cm for both management regions.

| Parameter | North | South |
|---|---|---|
| $L_\infty$ | 126.0 cm. | 129.2 cm. |
| K | 0.1080 | 0.1198 |
| $L'$ | 30 cm. | 30 cm. |

The standard deviation of the mean length (above $L'$) was used to develop a standardized normal distribution with mean 0 and standard deviation 1. The truncated distribution was rescaled so that unit area was obtained between the values of the standardized normal distribution corresponding to $L = L'$ and $L = L_\infty$ . The median of the resulting distribution and boundaries of 95% of the distribution were estimated conditional on given values of $L_\infty$, K and $L'$. The corresponding range in Z thus does not reflect variance contributed by error in estimation of $L_\infty$, K or $L'$, nor any covariance among terms. These estimates should be considered minimum estimates of the potential range in Z.

Estimates of Z by area and year, and minimum 95% confidence intervals are presented in Tables C43 and C44. SARC 31 recognized that if the assumption of M=0.2 is correct, the Beverton-Holt length-based method using L'=30 gives unreasonable estimates of $F_{threshold}$. However, the analysis showed an underlying trend in total mortality consistent with increasing landings and decreases in average and maximum size in survey time series, and the SARC considered the Beverton-Holt estimates as a useful index of trends in total mortality.

Mortality rates were estimated using Heinke's method from NEFSC bottom trawl survey abundance at age data (Table C45). The annual estimates are highly variable and many result in unreasonable estimates. This is probably due to inter-annual variations in catchability coupled with the overall low catch rates of goosefish in the NEFSC surveys.

Catch curve estimates of Z were calculated from the NEFSC winter survey by following the 1993-1995 cohort abundances over time (Figure C83). The estimates of total mortality (Z) ranged 0.29 - 0.40.

Catch curves were also fit to abundance at age data from the cooperative survey (Figure C84). The resulting estimates were Z=0.43 for both management regions and for the regions combined.

Exploitation ratios were calculated from the cooperative survey (Table C46). The estimates were produced using two methods: using landings and exploitable biomass from the cooperative survey (> 40 cm north, > 52 cm south), and using catch (landings plus discards) and total biomass from the cooperative survey. In each case, landings (catch) were added to the cooperative survey estimate of biomass to derive a proxy for biomass at the beginning of 2000, and the cooperative survey biomass was taken as biomass at the beginning of 2001. The exploitation ratio was calculated using the average between 2000 and 2001 biomass estimates. The estimates were produced under varying assumptions re. net efficiency and methods for estimating tow distance for the Mary K. This produced estimates of F ranging from 0.10 (north, low efficiency net assumption, total biomass method) to 0.43 (south, 100% net efficiency assumption, inclinometer data for Mary K, exploitable biomass method). Not surprisingly, the catch and biomass method produced lower estimates of F than the exploitable biomass method.

### Yield Per Recruit

In response to the SARC 31 research recommendation to re-evaluate reference points for goosefish, the Working Group developed an age-based yield per recruit analysis (Thompson-Bell model) to provide potential alternative reference points. Yield per recruit reference points (Fmax as a proxy for Fmsy, F0.1 as Ftarget) are suggested by the WG as potential alternatives to the current fishing mortality reference points, which have not proven to be very useful in practice. Another potential source of reference points and evaluation of current status is the Bayesian production model (below), for which reference points expressed on a ratio basis (F/Fmsy, B/Bmsy) are likely to prove more stable and reliable than absolute estimates of F, Fmsy, B and Bmsy.

Since the SARC 31 assessment, new information is available on age, growth, and maturity of goosefish from NEFSC research trawl surveys during 1992-2001 and the cooperative survey in 2001. Age, growth, and maturity data from NEFSC winter, spring and autumn surveys during 1992-2001, from the cooperative survey, and from the studies of Armstrong (1987; Georges Bank to Mid-Atlantic Bight) and Hartley (1995; Gulf of Maine) provided information on age and growth used for the yield per recruit analysis.

Mean weights at age for the catch and stock were based on age and individual fish weight data collected in NEFSC winter, spring, and autumn surveys during 1992-2001 (n = 3538 fish). Data were available for ages 0-10, for fish from 9 to 96 cm total length, and 0.01 to 14.08 kg. These data showed very similar patterns in length and weight at age as those from the Hartley (1995) study and the cooperative survey. Patterns in length and weight at age were very similar for fish in the northern and southern management areas in both the NEFSC surveys and the cooperative survey. Mean weights at age in the catch and stock for ages 11-15 were estimated from a Gomphertz regression based on NEFSC survey 1992-2001 individual fish mean weights at age (Table C47).

Maturity estimates from the cooperative survey were similar to those reported by Armstrong (1987) and Hartley (1995), with $L_{50}$ for female goosefish at 40 cm (age 4.7) in the northern area and 46 cm (age 5.1) in the southern management area. NEFSC survey data for 1992-2001 (n=3302) indicated an $L_{50}$ of 41.0 cm for females (age 4), 35.2 cm for males (age 3), and 37.7 cm (age 4) for combined sexes. Guided by this information, the analysis assumed no mature fish at ages 0 to 3, 50% maturity at age 4, and 100% maturity at ages 5 and older (Table C47).

Selection patterns were based on length frequencies of kept and discarded goosefish from sea sampling, length frequencies from port sampling, consideration of the NEFSC and cooperative survey length frequencies for

2001, and work performed for the SARC 31 assessment to estimate selection patterns for different components of the fishery (Table C14, Figures C6 and C7). Age 5 fish were considered nearly fully recruited to the fisheries (S = 0.90) and age 6 fish fully recruited (S = 1.0). Selection at ages 2-4 were roughly based on the "Trawl catch vs Winter Survey" selectivities at length provided in Table C14, with an upward adjustment to nominally account for some discarding at those ages. Ages 0-1 (fish < 20 cm) were assumed to have zero selection by the fisheries (Table C47).

Yield per recruit for the above combination of mean weights, maturities, selection at age, and natural mortality rate assumed= 0.2 provided estimates of F0.1 = 0.138, Fmax = 0.197, and F20% = 0.295 (Table C47).

Sensitivity of the analysis to alternative ages of knife edge recruitment to the fisheries indicated that significant gains in yield per recruit could be realized by increasing the age of entry to the fisheries (Figure C85). The partial selection pattern analysis (Table C47) provides a comparable maximum yield per recruit (0.93 kg/recruit) as knife-edged entry to the fisheries at age 3 to 4 (about 0.9 kg/recruit; Figure C85).

Using the partial selection pattern analysis (Table C47) as a starting point, yield per recruit was also examined under the assumption that discards cause mortality but do not contribute to landings. This was done by splitting the selection pattern into "landings" and "discard" components. The minimum size regulations in the northern (43 cm or 17 inches total length, age 3) and southern (53 cm or 21 inches total length, age 4) management regions were used to determine the proportion of catch at each age that would be discarded. In the north, all fish less than or equal to age 3, 90% of fish age 4, 40% of fish age 5, and a small percentage of ages 6 and 7 would be discarded. In the south this discard ogive was shifted one age older.

Explicitly accounting for discards causes Fmax to decrease from 0.197 (Table C47) to 0.187 in the north and 0.177 in the south. The associated yield per recruit also decreases from 0.931 (Table C47) to 0.890 in the north and 0.842 in the south (Figure C86). Given a fixed minimum size regulation, increasing the age at 50% selection causes increases in the landed yield per recruit (Figures C85 and C87).

## BAYESIAN SURPLUS PRODUCTION MODEL ANALYSES

The Southern Demersal Working Group developed surplus production models for northern-area, southern-area, and combined area monkfish using the most recent assessment data for review by the SARC. This work is an extension of the working paper "*A discard with catch error model of monkfish biomass dynamics*" presented at SARC 31. The primary differences in the new modeling approach compared to the approach documented at SARC 31 are:

- discard fractions are lower (assumed to be 10% of total catch weight) during 1964-1994 as suggested by the SARC 31 review
- a combined-area model is also developed to address the possibility that biomass dynamics are better approximated with a single population approach
- the surplus production curve may be right or left skewed (Pella-Thomlinson

production model) to account for the possibility that the stock is more or less resilient to harvest as biomass declines

- the revised model includes the swept-area biomass estimates from the cooperative survey as an index of total stock biomass with measurement error

Four surplus production models with similar underlying assumptions were initially developed. Each of the four models used the NEFSC autumn survey weight per tow index as a measure of relative population biomass to fit a Pella-Thomlinson surplus production model. The four models represented:

1. Northern stock area biomass dynamics during 1964-2000
2. Southern stock area biomass dynamics during 1967-2000
3. Combined area stock dynamics during 1964-2000
4. Combined area stock dynamics during 1964-2000 including another relative abundance index from the NEFSC sea scallop survey during 1982-1999

Together, these 4 models represented three different scenarios:

(i) a two stock scenario (models 1 and 2);
(ii) a one stock scenario where the fall groundfish survey measured relative abundance trends;
(iii) a one stock scenario where the fall groundfish and the scallop survey both measured relative abundance trends.

Each of the four models was fit using total catch (as adjusted for discard) and survey indices for the relevant stock area . A total of 60,000 MCMC samples were generated from the posterior distribution using two chains with different starting points and thinning the chains by 2 to remove autocorrelation. Of these, the first 5,000 - 10,000 samples were discarded to burn-in the model, e.g. remove dependence on the initial parameter values. The next 20,000 samples were used to evaluate the convergence of the MCMC algorithm for the key model parameters. The remaining 30,000 samples were also thinned by a factor of 2 to remove autocorrelation and these, along with the samples from the convergence check, were used to compute the posterior distribution of model parameters and associated outputs.

After reviewing the initial model diagnostics and results, the Southern Demersal Working Group recommended several changes to the model to improve consistency with expected stock dynamics and fishery trends. In particular, the SDWG recommended that any foreign landings of monkfish, as reported in the online NAFO statistical databases, should be included in the input catch time series. It was agreed that this could be done only for the combined-area monkfish models because there was no way to apportion the NAFO foreign catches to the appropriate northern or southern stock area. The SDWG also indicated that the assumption about catch errors due to misreporting or discarding were probably appropriate and recommended that these be included in the final model runs. The SDWG also considered the assumed discard fraction for 1964-1992 to be reasonable and recommended that this be applied to the domestic fishery landings totals. Similarly, the SDWG recommended using the observed fishery discard fractions for 1996-2000. The SDWG also indicated that it was most appropriate to incorporate the swept-area estimates of stock biomass in 2000 as an index of absolute stock biomass if possible.

Last, the SDWG recommended that the four baseline models be run for the time period 1980-2000 to provide a sensitivity analysis of the effects of excluding the earlier portion of the time series where some questions were raised about the accuracy of the reported catches.

All of the SDWG recommendations were addressed in the final model runs. Results of the final runs for the northern and southern stock areas are presented in Table C48. Each of the runs included the cooperative survey biomass estimate as an index of total biomass in the stock area using a multiplicative lognormal error term.

Results-Bayesian Surplus Production Model
Convergence diagnostics were the GR plots showing the ratio of model estimates of within chain variance to mixed-chain variance for key model parameters. In most cases the two ratios either approached unity or stayed within the interval of [½, 2]. This suggested that the chains were reasonably well-mixed, since the expected value of the variance ratio approaches unity in the limit as the chain length becomes very large. Given the large number of parameters in the model (80+ parameters/unobservables), this was considered to be very satisfactory convergence for the purpose of evaluating the relative trends in biomass and/or fishing mortality, e.g., biomass relative to the biomass that would produce maximum surplus production.

Estimates of the mean and quantiles of the posterior distributions of key model parameters and important outputs are listed in Table C48. There the variable BRATIO is the ratio of stock biomass in year 2000 to the biomass that would produce maximum surplus production. The variable HRATIO is the ratio of the harvest rate in year 2000 to the harvest rate that would produce maximum surplus production. The parameter K is the carrying capacity. The parameter M is the shape parameter for the production curve in the Pella-Thomlinson model. The variable B2001 is population biomass at the start of year 2001. The variable BMSP is the population biomass that would produce maximum surplus production (MSP). The variables qFALL and qSCALLOP are the catchability coefficients for the fall groundfish and the scallop survey biomass time series. The parameter r is the intrinsic growth rate of the stock. The parameter sigma2 is the process error variance, while the parameters tau2FALL and tau2SCALLOP are the observation error variances for the fall groundfish and the scallop survey biomass time series.

Model results indicated that fishing mortality has increased and stock biomass has decreased during the assessment time series of 1964-2000. Current stock biomass appears to be at or below BMSP. In particular, the median estimates of BRATIO for the northern and southern stock areas were 1.02 and 0.57, respectively. Current fishing mortality appears to be above HMSP. In particular, the median estimates of HRATIO for the northern and southern stock areas were 1.85 and 3.82, respectively. In addition, the SARC noted that the estimated production curve was right-skewed in each scenario; this indicated greater resilience to fishing pressure than would be expected under a Schaefer surplus production model.

The evaluation of monkfish status in relation to surplus production reference points for overfished condition and overfishing was conditional on which model scenario, e.g. scenarios (i) or (ii) or (iii), was considered to be most representative. The SARC did not reach a consensus as to which model scenario was most appropriate. However, the SARC did note that scenario (iii) had poor residual patterns for the relative abundance indices and that, under this scenario, the model predictions did not fit the observed data very well. Regardless, each of the model scenarios was consistent with the observed trends in the fall groundfish biomass time series which indicated a long-term decline in biomass. Similarly, each of the model scenarios showed an increasing trend in exploitation rate through time with peak values in the 1990s.

## EVALUATION OF STOCK STATUS WITH RESPECT TO REFERENCE POINTS

Northern Region

For SAW 23 and SAW 31, fishing mortality for goosefish was estimated from autumn survey length frequencies (NEFSC 1997; NEFSC 2000). This approach resulted in an unfeasible estimate of $F_{threshold}$ for the northern component. The analysis showed an underlying trend in total mortality consistent with increasing catches and decreases in average and maximum size but F could not be estimated reliably. Therefore, SARC 31 concluded that although current proxies are considered unreliable, the estimates of Z could be taken as a total mortality index, and concluded that overfishing was occurring. By these same criteria, the current assessment indicates that overfishing is occurring (Table C43, 1997-2000 average).

The current three-year moving average catch per tow (kg/tow from NEFSC offshore autumn research vessel survey) of 1.43 kg/tow is below the 33rd percentile of the 1963-1994 series, 1.460 kg/tow (Table C49), the biomass threshold below which the stock component is defined to be overfished. The moving average has been below the 33rd percentile since 1989, and is well below the biomass target of 2.496 kg/tow (median of three-year moving average during 1965-1981).

Southern Region

For SAW 23 and SAW 31, fishing mortality for goosefish was estimated from autumn survey length frequencies (NEFSC 1997; NEFSC 2000). The analysis showed an underlying trend in total mortality consistent with increasing catches and decreases in average and maximum size but point estimates of F were not considered reliable. Therefore, SARC 31 concluded that although current F proxies were considered unreliable, the estimates of Z could be taken as a total mortality index, and concluded that overfishing was occurring. By these same criteria, the current assessment indicates that overfishing is occurring (Table C44, 1997-2000 average).

The current three-year moving average catch per tow (kg/tow from NEFSC offshore autumn research vessel survey) of 0.427 is below the 33rd percentile of the 1963-1994 series of 0.750 kg/tow (Table C49), the biomass threshold below which the stock component is defined to be overfished. The moving average has been below the 33rd percentile since 1987, and is well below the biomass target of 1.848 kg/tow (median of three-year moving average during 1965-1981). The current three-year moving average biomass indices are also well below the proposed revised biomass target for the southern region of 1.846 kg per tow, and the proposed revised biomass threshold of 0.704 kg per tow (Table C49).

Trends in stock biomass, recruitment, and mortality

For the northern component, NEFSC autumn and spring research survey indices show an overall decline in biomass between 1984 and 1999; however, biomass indices in the north increased in 2000 (Tables C15 and C16, Figures C10 and C12). The increase in 2000 reflects increases in both spring and autumn survey abundance indices since 1998 (numbers per tow, Figures C11 and C13). The improved recruitment during the 1990s reflects contributions from the 1992, 1993, 1998 and 1999 year classes. However, the maximum and mean lengths of goosefish in survey catches (Figure C16) remain low.

For the southern component, the NEFSC spring and autumn surveys indicate that stock biomass and abundance have fluctuated around the time series low since the mid-1980s (Tables C20 and C21, Figures C21 and C23) . As for the northern component, decreases in the abundance of large fish in the spring and autumn surveys and decreases in the maximum and mean lengths of the survey catches suggest increasing fishing mortality rates over the time series (Figures C31 and C32). The NEFSC winter flatfish survey indicates no trend in biomass during the 1990s (Table C22, Figure C26); however, the survey has only been conducted since 1992.

For both stock components, indices of egg production (Figures C80-C82) mirror the progressive decline in abundance of larger fish in survey catches.

## SARC COMMENTS

The SARC discussed the basis for assessing goosefish as a single stock versus two stocks but did not feel sufficient information exists to make this biological determination. Information presented in favor of two stocks was the recruitment series and minimal adult migration while similar growth patterns and maturity schedules as well as a genetic study favored the one stock hypothesis. In the previous assessment, growth was thought to be different between the two areas, but the industry cooperative survey did not find a difference. It was noted that the genetic study did not provide definitive evidence because low rates of mixing could produce the appearance of a single stock when in fact there were two. Given that there is insufficient information to make the determination, it was decided that the two assessment units approach would be continued. In addition a combined unit is considered.

The SARC noted that the choice of number of management units for this species is independent of the number of assessment units. The use of two management units may be required because landings by gear type differ in the two current regions. Of special note is the apparent distinction between the proportion of landings coming from directed trips in the north versus south and the associated discarding implications of size regulations. In contrast, the use of a single management unit provides consistent regulations for all areas, reducing the complexity of management, but could potentially allow overfishing of one stock if in fact multiple stocks are contained in the management unit.

The SARC discussed potential alternatives for goosefish overfishing definitions because the method used to set the values, i.e. length based Z, has inherent flaws and $F_{threshold}$ in the north is implausibly low. Sufficient information now exists to estimate current fishing mortality rates by age and so yield per recruit analyses, perhaps using different natural mortality rates by sex, could be used to set the reference points. It was noted that the overfishing definition needs to be set in a metric that can be measured in the current year of an assessment to allow determination of current status. Consensus was reached that many lines of evidence point towards overfishing occurring in both the northern and southern units.

The SARC continues to support further development of the Bayesian surplus production model for goosefish assessment. Questions arose as to the appropriateness of the catch data for years 1964 to 1979 when landings are thought to be severely under-reported. However, truncating the time series used in the model to 1980-2000 resulted in unrealistic values for the intrinsic growth rate. Thus, while the SARC does not find a problem with the modeling approach, the data appear to be insufficient to support such modeling at this time.

The SARC commends the collaboration exhibited in the goosefish industry cooperative survey conducted in 2001. This cooperative venture produced new information on growth, maturity, distribution, cannibalism, catch rates, and selectivity that was directly applicable to this assessment.

## RESEARCH RECOMMENDATIONS

1) Research should be continued to define stock structure, including genetic studies, reproductive behavior analyses, morphometric studies, parasite studies, elemental analyses, and studies of egg and larvae transport.

2) The SARC recommends changing the overfishing definitions for goosefish. Research on yield per recruit for goosefish should examine the effect and possible causes of differential natural mortality rates by sex, methods to estimate gear selectivity, and the incorporation of discards.

3) Surplus production modeling should continue with special emphasis placed on uncertainty in under-reported catches and population size prior to 1980.

4) Size selectivity studies should be conducted in the trawl fishery to investigate the potential effectiveness of minimum mesh size and shape regulations to reduce discards of undersize monkfish. Additionally, comparative studies of the size selectivity and catchability of trawls and gill nets should be undertaken in order to understand the differences in the numbers of large fish captured in the two gear types.

5) Another cooperative survey for monkfish should be conducted in 2004.

6) Improved sampling rates (as observed in 2000-2001) for commercial landings should be maintained, which should eventually lead to an age-based assessment approach for this species.

7) Tagging studies should be considered as a basis to evaluate adult movement and rates of growth.

8) Spatial distribution of mature and immature fish and the potential effects of size limits on fishing behavior should be evaluated as a basis for advising on strategies to minimize catch and discard of immature fish.

9) Indices of abundance should be developed from industry “study fleets,” including coverage from outside the depth and spatial range of the NEFSC research surveys.

## LITERATURE CITED

Almeida, F. P., D.-L. Hartley, and J. Burnett. 1995. Length-weight relationships and sexual maturity of goosefish off the northeast coast of the United States. N. Am. J. Fish. Manage. 15:14-25.

Armstrong, M. P. 1987. Life history of the goosefish, *Lophius americanus*. M. Sc. Thesis, College of William and Mary in Virginia. 80 pp.

Armstrong, M. P., J. A. Musick, and J. A. Colvocoresses. 1992. Age, growth and reproduction of the goosefish, *Lophius americanus* (Pisces: Lophiiformes) U. S. Fishery Bulletin. 90:217-230.

Beverton, R. J. H. and S. J. Holt. A review of methods for estimating mortality rates in exploited fish populations, with special reference to sources of bias in catch sampling. Rapprts et Proces-Verbaux des Reunions 140: 67-83.

Caruso, J. H. 1983. The systematics and distribution of the lophiid angler fisher: II. Revision of the genera *Lophiomus* and *Lophius*. Copeia 1:11-30.

Chikarmane, H.M. , Kuzirian, A.M, Kozlowksi, R, Kuzirian, M. and Lee, T. 2000. Population genetic structure of the goosefish, *Lophius americanus.* Biol. Bull. 199: 227-228.

Cleveland, W. S. 1979. Robust locally weighted regression and smoothing scatterplots. J. Am. Stat. Assoc. 74(368): 829-836.

Cochran, W. G. 1977. Sampling techniques. Wiley. NY.

Conan, G. Y. and E. Wade. 1989. Geostatistical mapping and global estimation of harvestable resources in a fishery of northern shrimp (Pandalus borealis). Int. Coun. Explor. Sea C.M. 1989/D:1, 11p.

Hartley, D. 1995. The population biology of the goosefish, *Lophius americanus*, in the Gulf of Maine. M. Sc. Thesis, University of Massachusetts, Amherst. 142 p.

Martinez, C.M. 1999. The reproductive biology of the goosefish, *Lophius americanus*. M.S. Thesis, University of Rhode Island.

Myers R. A. and P. Pepin. 1990. The robustness of lognormal-based estimators of abundance. Biometrics 46:1185-1192.

NEFSC. 1991. Report of the twelfth northeast regional stock assessment workshop (12th SAW). NEFSC Ref. Doc. 91-03.

NEFSC. 1997. Report of 23rd Northeast Regional Stock Assessment Workshop(23rdSAW). NEFSC Reference Document 97-05.

NEFSC. 2000. 31st Northeast Regional Stock Assessment Workshop(31st SAW). NEFSC Reference Document 00-15.

Pennington, M. 1983. Efficient estimators of abundance, for fish and plankton surveys. Biometrics 39:281-286.

Pennington, M. 1986. Some statistical techniques for estimating abundance indices from trawl surveys. Fishery Bulletin US 84:519-525.

Rao J. N. K. and C. F. J. Wu. 1988. Resampling inference with complex survey data. J. Amer. Stat. Assoc. 83:231-241.

Sissenwine, M. P. and E. W. Bowman. 1977. Fishing power of two bottom trawls towed by research vessels off the northeast coast of the USA during day and night. ICES C.M. 1977:B30.

Sitter R. R. 1992. A resampling procedure for complex survey data. J. Amer. Stat. Assoc. 97:755-765.

Smith S. J. 1990. Use of statistical models for the estimation of abundance from groundfish trawl surveys. Can. J. Fish. Aquat. Sci. 47:894-903.

Smith S. J. 1996. Analysis of data from bottom trawl surveys, in H. Lassen, ed., "Assessment of groundfish stocks based on bottom trawl surveys" NAFO Scientific Council Studies, 28, pp 25-53.

Smith S. J. 1997. Bootstrap confidence limits for groundfish trawl survey estimates of mean abundance. Can. J. Fish. Aquat. Sci. 54:616-630.

Smith S. J. 2000. Notes for National Graduate School Fish Biology and Fisheries Course on: Design and analysis of trawl/dredge surveys for abundance. Sep. 20-25 1998, Tvarminne, Finland.

Smith S. J. and S. Gavaris. 1993. Improving the precision of fish abundance estimates of Eastern Scotian Shelf cod from bottom trawl surveys. N. Amer. J. Fish. Manage. 13:35-47.

Steimle, F.W., W.W. Morse, and D.L. Johnson. 1999. Essential fish habitat source document: goosefish, Lophius americanus, life history and habitat characteristics. NOAA Technical Memorandum NMFS-NE-127.