

## Appendix 1

### SNE/MA Yellowtail flounder Industry Meeting Participants: February 27<sup>th</sup>, 2012

<b>Name</b>	<b>Affiliation</b>
Larry Alade	NEFSC
Adam Barkley	SMAST
Gene Bergson	Harbor Blue Seafood
Jeff Bolton	Atlantic Capes Fisheries
Jason Boucher	SMAST
Katie Burchard	NEFSC
Steve Cadrin	SMAST
Richie Canastra	Buyers and Sellers Exchange NE
Peter Cura	F/V Fisherman
Dan Eilertsen	Nordic Inc
Ronnie Enoksen	Nordic Fisheries
Dan Georgianna	SMAST
Brian Gervalis	NEFSC
Dan Goethel	SMAST
Eric Hansen	F/V Endeavor
John Haran	Northeast Fisheries Sector 13
John Hoey	NEFSC
Robert Johnston	NEFSC
Jim Kendall	New Bedford Seafood Consultants
Chris Legault	NEFSC
Dave Martins	SMAST
Linda McCann	Northeast Fisheries Sector 7 & 8
Chris Medeiros	Quinn Fisheries
Cate O'Keefe	SMAST
Peg Parker	Commercial Fisheries Research Foundation
Ted Platz	Ocean Harvest
Charlie Quinn	Quinn Fisheries
Judith Rosellon	SMAST
Daniel Salerno	Northeast Fisheries Sector 5
Ron Smolowitz	Fisheries Survival Fund
Kevin Stokesbury	SMAST
Mark Terceiro	NEFSC
Doug Zemeckis	SMAST

**54th Northeast Regional Stock Assessment Workshop  
Southern New England/Mid Atlantic Yellowtail Flounder  
Pre-Assessment Meeting with Fishermen**

Monday February 27, 2012 10:00am  
School for Marine Science & Technology (SMAST)  
200 Mill Road  
Fairhaven, MA  
Room 158

Meeting Agenda:

- Welcome & Introductions
- Review of the 2008 stock assessment
- Growth, maturity and natural mortality
- Preliminary fishery data
- Preliminary survey data
- SMAST Industry-Based Survey
- Discard mortality
- Stock assessment models
- Discussion

Stock assessment scientists will review the most recent stock assessment of southern New England/Mid Atlantic yellowtail flounder, present updated information from the fishery and surveys, and summarize the plan to update the stock assessment this spring.

Steve Cadrin – School for Marine Science and Technology:

Opening introductions

Meeting agenda

Larry Alade – Northeast Fisheries Science Center:

Review of SAW 54 Terms of Reference:

1. estimate landings/discards
2. present survey data including vessel change
3. stock definition
4. estimate fishing mortality, recruitment, total and spawning stock biomass
5. describe causes of variability in annual recruitment
6. update Biological Reference Points
7. evaluate stock status with models
8. short-term projections and risk analysis

Timeline:

Data meeting: April 2 – 4, 2012

Model meeting: April 30 – May 4, 2012

SAW SARC 54 Review: June 5 – 9, 2012

Stock Status from GARM III (2008):

- Age 6+ VPA model formulation
- Natural mortality  $M=0.2$
- Assumed constant maturity at age
- Model years included 1973-2007
- $F_{MSY}$  proxy =  $F_{40\%}$
- Stock status = overfished ( $SSB = 3,508$ ) and overfishing occurring ( $F = 0.4129$ )

SAW 54 Updates/Inclusions:

- Re-evaluate all data sources and any data revisions
- Surveys: NEFSC Fall 1963-2010; NEFSC Spring 1968-2011; NEFSC Winter 1992-2007
- Survey calibrations applied to NEFSC Spring 2009-2011 and NEFSC Fall 2009-2010
- Revise landings and discards data based on database change in 2007
- Examine stratified discard estimate by area for scallop fishery, including analysis of observer coverage levels by area
- Include catch from scallop trawl vessels
- Include 2010 At-Sea Monitoring data
- Examine the discard mortality assumption (currently = 100%)
- Examine biological influences on recruitment – cold water pool indices
- Examine growth, maturity and stock structure assumptions

- 

Presentation of Preliminary Data for SAW 54:

- Fishery data (landings and discards)
- Effort data
- Survey data
- Survey distributions

**Discussion of presentation:**

- Industry has seen larger fish than observed in the surveys, are any of the methods in the survey flawed or biased?
- There has been a strong decline in stock level since the early 1970s
- There has been two decades of poor recruitment
- Why is the level of discards in the scallop fishery so much greater than landings?
- Fishery has not been landing yellowtail and majority of catch is discarded
- The fishery has largely been a discard fishery for the last 6-8 years due to trip limits
- Industry has observed larger fish in the Northeast (i.e., Georges Bank) and small fish in the Southwest (i.e., Mid-Atlantic)

Katie Burchard – Northeast Cooperative Research Program:

Utility of electronic logbook data for assessment

- NOAA Study Fleet coverage in Southern New England/Mid-Atlantic stock area 2007-2011
- More observed effort in Study Fleet in Statistical Areas 537,539,611 than observer coverage
- Study Fleet data can be used to verify and complement observer data
  - Self-reported data is accurate compared to observer data
  - Can be used as an additional data source in the assessment
  - Study Fleet vessel level data can be more accurate due to consistency in reporting by captains
  - Study Fleet data is less random than observer data

**Discussion of presentation:**

- Possibly include any scallop dredge Study Fleet data to verify discard data
- Industry wants to push the use of Study Fleet data in the assessment process due to large investment in data collection

- Long-term plan for Study Fleet would include a reduction of observer coverage and increase in level of self-collected data
- Important to note that industry-collected data can be used to verify observer data

Rob Johnston – Northeast Fisheries Science Center:

Comparison of Sweep Type for Survey Calibration

- Albatross replaced with Bigelow in 2007/2008
- Limited time for vessel calibrations
- Decision to change entire survey system with new vessel
  - New net
  - Potential use of 2 different sweeps in different areas
  - Timeline for testing too short
  - Result in broken time series
  - Less efficient roller sweep chosen for survey purposes

Studies conducted to examine sweep efficiency:

- Twin trawl with cookie sweep on one side and roller sweep on the other, separated by a box in the middle
- Paired trawl study with two vessels, one towing a cookie sweep and the other with roller sweep
- Goal: evaluate efficiency, size selectivity, fill in gaps in biological sampling
- Results:
  - No significant differences for catchability by season
  - No differences in size selectivity
  - Twin trawl experiment:
    - Cookie sweep and rock hopper sweep compared closely
    - Cookie sweep significantly more efficient, however with a catch rate approximated at 1.2 : 1
  - Paired trawl experiment:
    - Cookie sweep significantly more efficient
    - Result very different from twin trawl
    - Cookie sweep efficiency approximated at 2 : 1 over rock hopper sweep
    - Unknown vessel effects may explain results

•

**Discussion of presentation:**

- Was there ever a direct comparison between the Albatross and Bigelow with all of the parameters identical, then varied (including tow time, sweep choice, tow speed)?
- Many calibration tows were conducted, did not directly compare catch from Albatross with 30 minute tow to Bigelow with 20 minute tow
- Tow time has a strong influence on catch – 30 (Albatross) vs. 20 (Bigelow) minutes is a major change and could have further reduced the efficiency of the rock hopper sweep due to the herding behavior of flounder
- Twin trawl comparisons do not account for herding behavior. It is likely that there was a significant amount of crossover behavior from the fish and the results that show similar efficiency may not be accurate.
- The pair trawl experiment results showed that the cookie sweep was approximately 2 times more efficient than the rock hopper sweep. Vessel effect alone does not adequately explain the results.
- Trouser trawl experiments have shown similar bias in efficiency estimates as a twin trawl due to the herding behavior and net crossover.
- The survey sweep (rock hopper sweep) should be compared with the NEAMAP survey vessel, F/V Darana R.

Adam Barkley – School for Marine Science and Technology:

Yellowtail Flounder Industry-Based Survey

- Rhode Island DEM collaborated in an Industry-Based Survey of yellowtail flounder in Southern New England, including the Nantucket Lightship Closed Area in 2003-2005
- Results from the 2003-2005 IBS were used in the GARM III SNE/MA yellowtail assesment
- Results suggested no difference in abundance or biomass inside vs. outside of the Nantucket Lightship area, and less than 3% of the stock inside the closed area
- SMAST replicated the survey in the Fall of 2011 to determine if there have been changes in the spatial distribution of the stock and utilization of the closed area
- SMAST survey used same net and vessels
- Results from the survey showed more catch outside the closed area than inside
- 57% of fish caught inside closed area were sub-legal size
- Exploitable biomass was estimated at 1,042mt
- Results showed a change in % biomass in open vs. closed area since the 2005 survey

•

**Discussion of presentation:**

- Could the closed area be less productive due to the fallow bottom? Does continuous towing increase productivity due to increased food availability, reduced predators?
- Very high abundance of skates and dogfish in Southern New England could be causing increased natural mortality of flounder.

**The assessment could examine consumption rates of elasmobranchs**

- Clam boat effort has increased in Southern New England in the last decade and the effects of clamming on the seafloor could impact food availability.
- Are we sure that the current stock boundaries are correct? Historically there were clear differences in the fish in the eastern vs. western parts of the Nantucket Lightship Area, and extending north into the channel.
- Were the survey methods from 2003-2005 identical to the 2011 survey?
  - Tow time varied between survey: 2003-2005 survey focused on tow distance;
- 2011 survey set a tow time of 20 minutes

Adam Barkley – School for Marine Science and Technology:

Discard Mortality Estimation

- Reflex Action Mortality Predictor (RAMP) was tested on stressed and unstressed yellowtail flounder
- Process for testing included commercial capture, acclimation in test tank, branding for identification, exposure to stress through towing in trawl or held as a control in cages
- 7 RAMP tests conducted
- Factors affecting mortality include air exposure, tow time and stress from being towed
- Method was applied to yellowtail flounder caught in scallop dredges on Georges Bank, and trawl vessels in Southern New England
- Results show a discard mortality level of 82% for dredge-caught flounder and 81% for trawl-caught flounder

**Discussion of presentation:**

- This technique could be applied to skates in the gillnet fishery to examine discard mortality.

## Appendix 2

### SNE/MA Yellowtail flounder Data Meeting Participants: April 2-4, 2012

<u>Name</u>	<u>Organization</u>
Larry Alade	NEFSC
Adam Barkley	SMAST
Katie Burchard	NEFSC
Steve Cadrin	SMAST
Kiersten Curti	NEFSC
Greg DeCelles	SMAST
Brian Gervelis	NEFSC
Dan Goethel	SMAST
Jon Hare	NEFSC
Dvora Hart	NEFSC
Anne Hawkins	NEFMC
John Hoey	NEFSC
Chris Legault	NEFSC
Richard McBride	NEFSC
David McElroy	NEFSC
Murali Mood	NEFSC
Tom Nies	NEFMC
Paul Nitschke	NEFSC
Loretta O'Brien	NEFSC
Megan O'Conner	NEFSC
Cate O'Keefe	SMAST
Mike Palmer	NEFSC
Greg Power	NERO
Dave Richardson	NEFSC
Eric Robillard	NEFSC
Gary Shepherd	NEFSC
Ron Smolowitz	Coonamessett Farm
Katherine Sosebee	NEFSC
Mark Terceiro	NEFSC
Michele Traver	NEFSC
Susan Wigley	NEFSC
Tony Wood	NEFSC

## **SNE Yellowtail Data Meeting Notes: April 2-4, 2012**

### **WG Consensus**

- No evidence for change in stock structure for this assessment
- Adopt the proposed base (time series) and alternative (5-year moving average) as options for observed maturity proportions
- Larval index may be useful as SSB index for model calibration
- Use the NEFSC Survey-based L-W relationship for 1994 and later years
- Adopt an alternative lifetime  $M = 0.3$  and to scale the Lorenzen curve to age 9 with spring, fall and commercial ages pooled. This is likely to be the preferred alternative with a sensitivity of constant 0.2 and 0.3 across all ages. Given that natural mortality estimates range from 0.3-0.5 and this stock has experienced high fishing mortality over the time series, WG consensus is that a lifetime  $M$  of 0.3 is reasonable.
- Information on the cold pool index should be incorporated into the discussion of the vulnerability TOR.
- Given that 85% seems to be a lower bound on the RAMP-based discard mortality and some mortality likely occurs post-release, the WG agreed to use a value of 90% for commercial fishery discard mortality in the assessment.

### **WG Research Recommendations**

- Consider using fine-level stratification to develop discard estimates for scallop rotational areas, especially the Nantucket Lightship Area (NLS), for 2000 and later years.
- Develop approaches (e.g., hindcast ratios) to develop discard estimates for fishery strata with no observer coverage
- Update the length-weight parameters used to convert commercial landings (in weight) into numbers of fish. This could be accomplished by expanding existing data collection programs (e.g., Cooperative Research, Industry Based Surveys, NEFSC port sampling) to collect individual fish weights while collecting length and age data. This research recommendation is applicable to numerous species/stocks in the northeast, not just SNE/MA yellowtail flounder.
- The work on the influence of the cold pool and associated environmental parameters on yellowtail population dynamics has not been fully developed, and merits further research.
- If the volume of commercial landings increases in the future, ensure that adequate samples of the landings are obtained for all market categories on at least a quarterly basis.

### ***Daily Notes April 2 morning***

#### **Stock structure**

- Cadrin: Brown coined the term “Harvest stocks” – even if there is exchange between stocks, we need to manage separately if they respond differentially to harvest. It seems like recruitment dynamics are different among stocks; Phenotypic boundary likely driven

by temperature; Boundary between SNE and GB appears to be “squishy” and dependent on stock size

- Hare: Summary: Current stock definitions are appropriate, but we need to begin considering the northward shift in distribution documented in Nye et al. Is this a consequence of a shift in distribution or a difference in productivity? Currently unable to disentangle these two hypotheses; Greater differences in growth/maturity among stock areas earlier in the time series compared to later in the time series; Two hypotheses: 1) growth conditions becoming more similar among stock areas, or 2) greater mixing among stock areas
- Loretta: Did Jon Hare consider temperature changes when looking at changes in growth?  
Jon Hare: not yet.
- Cadrin: Trying to recall Friedland paper: Friedland found different growth patterns between GB and SNE, but found that growth differences became less pronounced. Friedland inferred greater mixing among areas; Cadrin feels that paper confirms vagueness of GB/SNE boundary, not increased mixing
- Legault: Stratum 16 becoming more dominant in terms of proportion of YT total catch. But 16 is in closed area 2 --- so differences could be due to management as well
- Megan: There are distribution differences by age, but some truncation of age-structure

**WG consensus: No evidence for change in stock structure for this assessment.**

**Maturity**

- McBride: Not collecting age-1 fish. Not sure if reason is because age-1's are not selected by the survey, or because all age-1 are males. Larry thinks it is likely selectivity.
- Cadrin: Age-1's in the spring are very small; therefore, not really caught in the spring survey, when maturity analyses are conducted
- Loretta: Are there two sets of eggs in the gonads?; McBride: in the spring, there is an unyolked set and a cohort developing for the current year. Also repeated batches through the summer. Would be unusual to have spawning before age-2
- Cadrin: Seems that the few fish that were called resting but histologically were immature do not impact maturity ogive. May be more appropriate to report proportion mature at age-2 --- Could then demonstrate insensitivity.
- Maturity: Sample issue leads to sample size issue in maturity which impacts curve fit.
- Maturity trends: Should we update the time series, or should we use some type of moving average to capture trends?
- Nitschke: Proportion mature of age-2 increasing, but A50 plots flat or even decreasing;
- Larry: But much variability around A50 (model estimate)
- Cadrin: But one is slope (A50) and one is position
- Cadrin: Assumed proportion of Age-2 mature could have big impacts on SSB

- Loretta: Since spawning season is in the summer, could we construct a maturity ogive in the fall to see if it further informs our analysis? Would also have more age-1's. We have a bit of a unique situation with spawning in the summer
- Loretta: If we are going to use annual weights, we should try to capture some temporal variability in maturity; If use a moving average, do not have an issue with time blocks. Suggests 5-year moving average
- Currently using a time-series average for maturity (age-2 would be most influential age)
- Richardson: Is dip in maturity in recent years due to selectivity changes with the Bigelow? If 1) larger age-2 individuals are the ones mature, and 2) the Bigelow is catching smaller fish, would the observed dip be due to selectivity?
- McElroy: Samples by age, by year --- collecting more age-1 in last three years....
- McBride: At least partly due to increased sampling in recent years
- Terceiro: Looking back in time, many age-1 samples in late 70's – early 80's. Therefore, at least partly due to stock size
- Cadrin: Maturity trends seem to be somewhat lagged with biomass – supports a density-dependent aspect of maturity
- Legault: Agrees with idea of using a moving average, but questions whether we have enough samples to use a 5-year moving average. Sample size is very limited in some years (2003-2008 at age-2), which would yield very imprecise estimates
- McBride: Could you just plot only those years with greater than X number of samples?
- Alade: The assessment traditionally uses the time-series average of the observed proportions-at-age
- Loretta, in cod: Fit annual curves to 5-year moving averages
- Legault: For the other YT stocks, it is difficult to fit a logistic curve to a single age. The logistic has two parameters, but we only have one piece of information for YT: Proportion mature at age-2.
- Terceiro proposes either a 1) 5-year average of observed proportions, or 2) time-series average of observed proportions
- Hare: Is there a size-correction for the last few years to account for the Bigelow?
- Terceiro: Is there a strong case for going against precedent?
- Loretta: Concern is that we may lose some dynamics by using time-series average.
- Cadrin: There may be some small age-2's that might now be sampled by the Bigelow but were not sampled by the Albatross. Provides support for the base-case
- Terceiro: But we did catch age-1 fish when stock size was much greater
- Terceiro: Base case = updated time-series average; Alternative case = 5-year moving average of observed proportions; Determine impact on SSB.

**WG consensus: adopt the proposed base (time series) and alternative (5-year moving average) as options for observed maturity proportions.**

### **Fecundity**

- Gary: Were you able to look at any fish post-spawning to account for attrition? Realized vs potential fecundity
- McBride: Cod equals <5%
- Terceiro: Take-home point: SNE most fecund of the YT stocks;
- Cadrin: Most dominant year classes were from low-stock sizes

### **Length-weight relationship**

- Larry proposes using 1) the most up-to-date data available (stock-specific estimates) for 1994-2011, and 2) the Lux relationship for pre-1994. Will apply spring for Jan-June, and fall for July-Dec
- McBride: Samples could be biased if only sampled during one portion of spawning season
- Wigley: Differences between commercial and survey length samples? Is it more appropriate to use survey relationships for discards but commercial relationships for landings?
- McElroy: If timing of spawning shifts and survey timing is constant, could impact length-weight relationships.
- Cadrin: Is torn regarding best way forward; Recommends looking at sample sizes from Lux and current analyses; Lux had very few fish smaller than 25 in the spring; Had quite a few small fish in the fall → Similar to survey, age-1's showing up in the fishery in the fall, but not in the spring
- Commercial catch-at-age: Not many age-1's post-1994.....
- Reserving judgment until see differences in sample size between studies; also need to decide whether to use survey length-weights for discards and fishery length-weights for landings.

### **Larval index**

- Nitschke: Did the two peaks line up with the two big assessment year classes? Dave doesn't think of it as an index of recruitment
- Cadrin: Sullivan et al attributed year class success to settlement success -- -therefore, could have high larval index but not high recruitment.
- Hare: Larval index is generally viewed as an index of SSB, not recruitment
- Legault: Is there an estimate of the variance? Richardson: Tim Miller can calculate the CV's using an MLE approach.

- Terceiro: Will need some type of precision estimate for input into a statistical catch-at-age model

**WG consensus is that larval index may be useful as SSB index for model calibration, Dave R. will talk to Tim Miller about calculating CV's.**

#### **Returning to Length-Weight relationship**

- SNE sample size: ~ 3300 fish
- Lux: spring 418, Fall = 930; Size distribution: has very few fish less than 20 cm or greater than 45 cm in any season.
- Current study: Broader length distribution, increased sample size, more recent study

**WG consensus is to adopt Larry's recommendation to use the NEFSC Survey-based L-W relationship for 1994 and later years.**

#### **Natural mortality**

- Cadrin: Is this something that we estimate by species or by stock? We see older fish on Georges Bank; Terceiro: We are considering YTFI at large
- Greg: Are there any empirical estimates from tagging studies?
- Tony: Not directly on M --- the estimates that Tony recently derived were unreliable and ~ 1.6
- Gary: We are trying to look at the maximum age of the population; with the length approach, we are trying to predict the average maximum age (as opposed to picking the one extreme value and assuming it is representative of the population).

#### **April 2 Afternoon**

##### **Natural mortality**

- The group discussed retaining the currently assumed natural mortality rate of 0.2. The Lorenzen method suggests that for older ages this assumption may adequate, but neither the survey nor the commercial fishery catch a lot of older fish. The traditional  $3/T_{max}$  approach would lead to a higher M of 0.27 (given observed max age of 11 years), while other methods estimate 0.3 - 0.5. **The working group agreed on an alternative lifetime  $M = 0.3$  and to scale the Lorenzen curve to age 9 with spring, fall and commercial ages pooled.** This is likely to be the preferred alternative with a sensitivity of constant 0.2 and 0.3 across all ages. The WG discussed changing M over time, but while there has been some age truncation over time, it does not warrant a change in M.

**WG Consensus: Given that natural mortality estimates range from 0.3-0.5 and this stock has experienced high fishing mortality over the time series, WG consensus is that a lifetime M of 0.3 is reasonable.**

### **Cold Pool Index**

There is a link between geographic location, the extent of the Mid-Atlantic cold pool and the recruitment process. The cold pool is the preferred thermal habitat for YOY yellowtail flounder. When the cold pool is small there is less suitable habitat for settlement, while there is more suitable habitat when it is large. The temperature effect is significant, but explains less than half of the variance. In particular, the 1980 and 1987 year classes are not explained by the cold pool or spawning stock biomass. Yellowtail flounder settle in coldest part of cold pool. The WG suggested examining the center of the SSB using the larval data and whether it is closer to cold pool during these 2 years. The WG also suggested examining the scallop survey data for recruitment index. Information on the cold pool index should be incorporated into the discussion of the vulnerability TOR.

### **Discard Mortality Rate**

The WG discussed the duration of the SMAST RAMP and discard mortality study. The fish were kept up to 60 days, but the analyses used 20 days since most of the mortality occurred within this time frame. There were also controls in cages on the sea floor which had a lower ramp score. The tow times of 1-2 hours were approximately commercial tow times gave the fish a range of stresses. For the relationship between RAMP and mortality, only a range of values was needed before sampling the commercial activities. There was no direct evidence of additional mortality from predators or starvation, but there is likely some additional mortality. The fish with the lowest RAMP would be the ones more likely to evade predators. Commercial trips occurred in the Gulf of Maine (otter trawl) and on Georges Bank (scallop). The full range of temperatures is that occur throughout the year is likely covered for scallop dredge and more otter trawl trips are planned. Information on species composition and catch size is being collected and will be examined. Tow time does not seem to be a significant factor while air exposure is significant. There do not seem to be any size dependent differences in mortality. The WG discussed the types of discarding practices that have been observed. Some use shovels and picks, which likely increase mortality more than a conveyor system. There does seem to be consistency in discard mortality estimates (80-85% mortality) regardless of method. When fish are being caught for tagging, the tow times are short and the handling very different than on a regular commercial trip. For yellowtail flounder there have been few, if any, multiple releases by commercial fisheries.

**Prior studies by MA DMF suggest 33-50% mortality. Given that 85% seems to be a lower bound on the RAMP-based discard mortality and some mortality likely occurs post-release, the WG agreed to use a value of 90% for commercial fishery discard mortality in the assessment.**

### **Study Fleet Discard Estimation**

- There is likely more of a mix of types of trips in the NEFOP than in the Study Fleet. Discard rates in NEFOP are generally higher for large mesh otter trawl, but discards are estimated higher in Study Fleet. This needs to be checked. There is potential for use of these data as we now use At-Sea Monitor trips, but more exploration is needed. These data could be a good supplement to Observer program to fill in gaps in the coverage.

There is also potential to use the information for a CPUE index fleet. The difference between NEFOP and Study Fleet estimates of discards by species gets smaller as the amount of discards gets larger. The observer could be getting the estimate from the Captain.

### **Discard Estimation**

- The high values early in the time series are explained by few trips in some cells and also require imputation. The blended method seems reasonable based on the number of trips by region, CVs and the early high values. The small mesh otter trawl values in the late 1990s are driving the high cvs. The WG discussed the stratification used and whether the scallop dredge fishery should be stratified into open/closed access areas. For 2000 and 2002, there was differential observer coverage between open/closed areas with most of the coverage in the closed areas, which tend to have lower bycatch rates. The observer data are easily separated into open/closed areas, but the landings for expanding to total discards require additional work.
- For the purposes of stock assessment, the working group decided to use the GARM III approach for years prior to 2002 and use the SNE/MA stratification for 2002-2011. The SNE/MA stratification should be re-done with areas 611-613 included in SNE.
- Scallop landings from trawl gear are 2 types, landings on flatfish trips should be in with all trawls. Directed scallop landings with a scallop trawl (052). These have been separated and a decision on what to do prior to 2004 will have to be made.

### ***April 3 morning***

#### **Ageing QA/QC**

Eric Robillard and Sarah Emery were in attendance to discuss the QA/QT of ageing SNEMAYT. Steve Cadrin requested to see the validation study that was done, as well as reference the workshop that was attended regarding the ageing. Rich McBride suggested that poster that was presented at AFS by Larry and Sam also be used since it is a wealth of information.

#### **Discard Estimation**

Dvora Hart made a presentation on the Scallop Fleet discards. She suggested we use:  $T = (D/K_{\text{trawl}})/(D/K_{\text{dredge}})$ . This is because she feels we need to patch the years with no observer coverage. Currently, when we lack observer coverage, we look at the percent discards and apply a ratio.

#### **RESEARCH RECOMMENDATION: when looking at this issue, a more complex procedure should be considered other than apply a ratio.**

- Discard estimates used in the assessment and ACL monitoring should be consistent. It may help release the current constraints. We have done that for the fleet, but we still need the patch the years that have no data. It would be helpful to have more communication between the NEFSC and the RO.

- Tom Nies asked what the results would be if the areas were “open north” and “open south”. It would be a reasonable option to modify current stratification scheme to areas south of Long Island. Larry will run analyses with Dvora’s idea (develop alternative set of estimates in redefined areas for both trawl and scallop dredge). It will be a matter of looking at the current stratification vs the proposed one before moving forward. However, we cannot use it back in time. Before 2003, the coverage varies by year so we’d have to pool it, but from 2003 on, there was lots of observer coverage. Cadrin proposed that we use Larry’s current way prior to 2003 and Dvora’s way post 2003, but no decision will be made until we have a chance to look at the results.

### **Industry Based Survey**

- Greg DeCelles (SMAST) presented the Industry Based Survey (IBS) results. There was some discussion about the age frequency in the areas sampled. Age Age-1 total biomass is based on the length frequency and there is a lot of overlap in the Age-2’s.
- A member of the audience asked about the areas that were not able to be sampled due to the bottom. Yes, they are included in the biomass estimates. The RI and SMAST surveys are comparable. Both surveys encountered the same issue with sea bottom. There are some holes due to the amount of dogfish. Greg et al tried to compare apples to apples and keep the same spatial density. It was requested that Greg et al take select survey strata to get swept area for their 3 data points to compare. The age-length keys are available.
- Rob Johnston was able to maximize the comparison between the RI and SMAST surveys. It was suggested to get the Confidence Intervals from our survey, then add it to theirs. However, there is no replication between cells.

### **Commercial Landings**

- Larry presented landings information. The relative differences are fairly big due to the re-running of the analysis and updated length-weight relationship. It could also be from the imputing of the age-length keys. Yes, the AA tables were used. In the length frequencies, the mediums are not used (they are less than 10%); they assume the length frequency of the aggregate.
- No comprehensive age-length data are available from the 1950s. M is based on what we see in contemporary samples. The Royce paper has age compositions from the 1940s-1950s; few fish older than Age-6. The paper says that it is based on the environment, not necessarily all fishing. Spatial distribution can be part of the change, but it is definitely different from then until now. Steve Cadrin will write up a paragraph based on the Royce paper as to what supported those landings. It needs to be available to the SARC.

- It was suggested to use ASAP to plot the age compositions. Please plot proportion at age. It will give another interpretation. There was a clarification on how the z-scores were calculated. Larry needs to check the math on this one and re-do.
- It was requested that Larry make a table with the number of samples, possibly by quarter if there were enough to do it that way.

*April 3, 2012 Afternoon*

**Miscellaneous Discussion**

- Regulations: basically two broad stanzas of selectivity, up to the mid-nineties with no mesh size regs, then through the present; constantly are changing mesh size regs from then on.
- Ages, lengths and commercial length frequencies: Table of sample sizes - check the length numbers and age numbers. Are there some categories that are commercial and survey combined? Are the “unclassified” lengths stable over the years?
- In 1999 the assessment was rejected as the age and length sampling was so sparse. If you use an ASAP model do not use certain years where the sampling is poor, especially where there are samples for only one half of the year as the growth is not constant through the year.
- Length-weight relationships: for commercial, some of them are 50 years old and need to be updated. Observer coverage is pretty high, there should be some data collected by them, or the port samplers, or cooperative research project participants, need individual kept lengths and weights to improve the models. **WG recommends Research Recommendation on this issue.**
- Proportion mature at age 2: The best estimates of proportion mature at age 2 might be different between the Albatross years and Bigelow years because the Bigelow catches smaller fish and smaller mature age 2s will be caught.

### Appendix 3

#### SNE/MA Yellowtail flounder Model Meeting Participants: April 30 - May 2, 2012

<u>Name</u>	<u>Organization</u>
Larry Alade	NEFSC
Adam Barkley	SMAST
Liz Brooks	NEFSC
Katie Burchard	NEFSC
Steve Cadrin	SMAST
Jon Hare	NEFSC
Fiona Hogan	NEFMC
Chris Legault	NEFSC
Tom Nies	NEFMC
Paul Nitschke	NEFSC
Robert O'Boyle	Beta Scientific Consulting Inc.
Dave Richardson	NEFSC
Gary Shepherd	NEFSC
Katherine Sosebee	NEFSC
Mark Terceiro	NEFSC
Michele Traver	NEFSC
Susan Wigley	NEFSC
James Weinberg	NEFSC
Tony Wood	NEFSC

## **SNE Yellowtail Model Meeting Notes: April 30 – May 2, 2012**

### ***Daily Notes***

#### ***April 30 morning***

- The working group noted that there was a large increase in age 1 commercial catch which was likely driven more by revisions of the age-length key than by new discard estimates. This is because the discard estimates between GARM III and this assessment are similar. It may be useful to look at the ALK before the SARC. However, the number is not out of line with catches prior to 1994.
- The working group discussed whether to include the southern strata in the winter survey. The abundance of yellowtail flounder in those southern strata was high in the 1970s but by the late 1980s and early 1990s, yellowtail had disappeared from those strata. Therefore, it seems reasonable to exclude them from an index that began in 1992.
- The larval index was discussed by the working group. It was noted that the 2010 and 2011 indices increased significantly. The index was presented as a split series using Dave's method and as a single index using Tim's maximum likelihood method. There was a different mesh size used prior to 1987, and Tim's method attempts to account for the difference in selectivity. There have been comparison tows, but more are needed and work is underway to complete these comparative tows.

#### **VPA**

- For the VPA runs that end in 2008, the last year of spring survey age composition residuals are all positive. The working group discussed using the spring survey weights for SSB and catch. The group decided that these shouldn't be used for catch since the numbers are not scaled to total weight properly if catch weights are different. This has no impact on the fitting of the model. Since most of fishery occurs in the second half of the year, it would not be appropriate to use the spring survey weights at age for catch.
- The impact of different discard mortality rates was examined. The estimates of recruitment are not impacted by using 80, 90 or 100 percent discard mortality. The retrospective for F gets better with lower mortality.
- The retrospective for F gets worse with updates of the data so models with  $M=0.2$  and a lifetime  $M$  scaled to 0.2 were run. The retrospective for F was decreased, but the retrospective for SSB increased but was still low.
- The working group discussed the possible models, Run 15b (Lorenzen 0.3) and Run 16b (Lorenzen 0.2). Since the working group agreed to use and  $M$  of 0.3, run 15b should be the starting point.

- All model runs have no information in year T+1 since the spring 2012 survey is not finished yet. It is the same formulation from GARM III, but GARM III was in August and had the spring survey information for year T+1. The working group discussed lagging the fall survey forward a year and an age to get some information for year T+1, but decided that this formulation is closer to any ASAP configuration.
- The weights-ate-age used to derive the Lorenzen scaled M had an abrupt shift in 1994 so the M at age 1 shifts as well. The working group decided to use a time series average Lorenzen M scaled to 0.3.
- The working group picked a base VPA (Run 20) with time series Lorenzen M scaled to a lifetime M of 0.3. There is no patterning in the residuals and no indication of doming in the survey catchabilities. The winter survey qs are high but with the ground gear on the winter survey net, herding is expected between the doors and the net. The CVs on age 2 estimates in the terminal year are high but given that there is no spring survey estimate for 2012 they are not unexpected.
- The RI IBS in 2004/2005 and IBS in 2011 are less than mean biomass estimates so there are no apparent catchability issues. The retrospective pattern is underestimating fishing mortality in the terminal year. SSB at the start of the model was 24,000 mt, declined to lower levels and had two excursions to higher SSBs due to two large year classes. Recruitment has been poor since the 1987 year class although SSB is now starting to increase due to low F.
- The working group decided to use the average of 2006-2010 for selectivity and 2007-2011 for mean weights (2007-2011) for reference points and projections. Recruitment will be handled with 2-stanzas of empirical estimates split at SSBs of around 5000 (Rago will re-run the razor).

### ***April 30 Afternoon***

Working session – no meeting

### ***May 1 morning***

Work during the morning session compared the different ASAP models and decided whether or not to continue to the VPA or move forward to the ASAP model.

- Run 1 vs Run 2:
  - Run 2 broke up residuals a bit. Small improvement seen. Coincides with major changes from 1994 onward in the management regime.
- Why doesn't the VPA F trend follow that of ASAP? Because there are fixed blocks and they are different models. Multinomial model used for age compositions.
  - Winter survey q was about 2 in VPA, about 3 here.
  - The F-report is different than VPA but they both have  $M = 0.2$ . (VPA selectivity changes every year so be careful when comparing to ASAP.)
  - In the CV plot, there are occasional spikes due to the lack of sample data.
- Bob O'Boyle asked to compare partial recruitment between VPA and ASAP.
  - VPA Run 20 with  $M = 0.3$  compared to ASAP Run 16 to address Bob's request.
  - Recruitment patterns seem to match fairly well across all ages and are configured the same way; they are virtually identical.
  - F pattern general trend is very similar between the VPA and ASAP. ASAP is slightly smoother in later years.
  - Are there fishing effort trends that corroborate with the F trend? There's an increase in survey indices but a decrease in catch. There are 2-for-1 counting days at sea; the fleet has been trying to get fishing off SNEMA YT.
  - The shifts that are seen can be due to selectivity blocks (there are 6).
  - SSB patterns are similar between the two models; it is a little flatter with the VPA.
- Side by side comparison between VPA Run 20 and ASAP Run 16 to decide which model to use for the assessment.
  - The VPA shows recruitment to the fishery to be more gradual. ASAP shows full recruitment (95%) into the fishery at age 2 in the early time blocks.
  - There are 6 selectivity blocks, 1 fleet.
  - Bob O'Boyle requested to see differences between the two models over age and time. Chris Legault did this. ASAP F – VPA F (by age) and plotted.
    - Ages 4-6 are equally selected for both models.
    - No real strong patterns (Bob would like indices emailed to him.)
    - Last 10 years are very consistent between the two models.
    - Age 3 has more differences than ages 1 and 2.
    - Ages 4-6 are the same.

- Blocks are split: 5-6 are similar, 4 has lower selectivity at age 3. It doesn't shift as nicely as hoped, but the blocks are short. Some only have 3 years for estimating 3 parameters.
- Retrospective patterns for ASAP Run 16 (looking at the various diagnostics).
  - F- 2004 is a “high flier.”
  - SSB – ASAP is more consistent in direction (6 above, 1 below). If the two fliers are thrown out, it looks reasonable. The two fliers almost cancel each other out.
  - Recruitment in the last year is not well estimated. Both models have positives and negatives; they both bounce around.
  - Has Larry looked at the historical retrospective patterns yet? No because the beginning VPA is locked in.
    - Larry did a comparison using the GARM III VPA to “new” VPA (Run 20) to ASAP Run 16 for Jim Weinberg’s request.
    - Recruitment is scaled up by increasing M (as expected).
    - Average F is nearly identical between old and new VPA. ASAP handles F differently. Trend is basically the same in all 3 models.
    - SSB – still end up in the same place in 2011, regardless of the model.
- The SARC has given guidance to move to a statistical catch at age model. What are the panel’s thoughts?
  - Steve Cadrin says to use ASAP because there is more flexibility to improve the model.
  - Chris Legault says that it gives confidence in both models because they are both similar.
  - WG conclusion was to develop ASAP through the reference points and continue with ASAP model as the preferred model framework. Still need to decide how many selectivity blocks to use.
- ASAP Runs 17-19 are using the larval index.
  - Run 17 was taken out because it was agreed to use  $M = 0.3$  and that one uses  $M = 0.47$ .
  - Run 19 uses  $M = 0.3$ , splits are 77-87 and 88-11.
  - Larval index used as an index of SSB.
    - What happens when a split in the larval index isn't used (Run 20)?
    - There is substantial impact on SSB, with a large increase at the end of the series.
    - RMSE is very large, indicating a need to increase the input CV.
    - The residuals are strongly patterned.

- Run 19 is not used for comparison. You have to increase the CV and decrease the influence. Create a Run 21 to replace Run 20 (make CV = 0.3, effectively doubling the original CV).
  - CV = 0.3, use the comparison tool. With this CV, it allowed the fit to be closer to Run 16.
  - Liz Brooks suggested that if there are year specific CVs, to double them instead of using a constant. The original CVs are very close (0.13-0.15). Larry ran it with doubling the CV.
- Run 22 will use a different larval index calculation.

**May 1 Afternoon**

Tuesday Afternoon

- The ASAP model has a better fit with Dave’s larval indices than Tim’s model-based estimates. The retrospective is improved with the addition of the larval indices compared to without, so the working group decided to include the larval indices from Dave.
- The working group examined models with varying selectivity blocks. The 6 selectivity blocks seem to produce selectivity estimates that do not necessarily agree with the expectations from the regulations. However, the improvement to the model fit is enough to warrant keeping in all six blocks. The retrospective pattern is also reduced with 6 blocks, so the WG chose the 6 block model. The final model increased the CVs on the survey indices by 0.1 to reduce the mean-square residuals.

RunID	Selex	Blocks	Change in Parameters	Obj Function
22	6			4683
23	4	-6		4703
24	3	-9		4715
25	3	(+0.1 to sv cvs)	-9	4675
26	6	(+0.1 to sv cvs)		4640
27	5	(+0.1 to sv cvs)	-3	4652

- The working group reviewed an analysis by Steve Cadrin of SMAST of different Fmsy proxies. The stock was able to replace itself at F40 in both early and late years, but at F30 the stock would not have been able to replace itself in the later years using ASAP and VPA results. The working group concluded that F40 is a good proxy for Fmsy.

### ***May 2 morning***

#### SRFit VPA run 20

- No Ricker has been attempted because of work done back at the GARM suggesting this relationship was not reasonable for YT flounder.

#### Bootstrap outputs VPA run 20, AgePro VPA run 20

- Used Paul Rago's updated cut point (~4,000 mt), stock in 2011 is just under the breakpoint.

#### SRFit for ASAP run 26

- Everything the same except as for the VPA Run 20 except the fishery selectivity, with ASAP indicating a slightly higher fishery selectivity.

#### MCMC results; YPR

- F40% estimates from VPA and ASPA both about 0.3

#### Revisiting TORs

- Prepare plots that go back to SARC 36 for the historical retrospective: F, SSB, recruitment.
- WG chair noted that performance of the projections is NOT a term of reference for this assessment.
- The SR functions did not provide a good basis for BRPs. Steve Cadrin's work suggests F40 is an appropriate proxy. ASAP is the preferred assessment model.
- The WG noted that management in the near future is going to be about rebuilding. Long term SSBs at F40 are in the same neighborhood as what was being returned from the B-H S-R function.
- WG recommended projections with the existing and new reference points to beyond the rebuild year of 2014 to evaluate when the stock might be rebuilt under different BRPs and recruitment scenarios.

## *May 2 afternoon*

### Projections

- The WG should note the concern in the report regarding the likelihood that recruits will jump up a bin in the rebuilding scenario.

### Coldpool S-R- model

- Took run 26 and used modified ASAP which allows covariates in the S-R relationship to look at coldpool index. As the coldpool index goes down you have a higher predicted recruitment. Gives intermediate results between F40 run and the post-1990 recent low-recruitment scenario.

### TORs

- TOR8: Projection with recruitment since 1990 is most realistic? Are we in a new productivity regime that will last for the foreseeable future?
- Two aspects that may not be independent: the first is climatic warming and the second is the change in geographic range. We no longer have the geographic range of the stock that was associated with the large recruitments of the 1970s and 1980s; starting the recruitment in 1990 is a reasonable alternative. Putting it forward as a scenario to the SARC reviewers will be informative.

### Research Recommendations

No new model-related research recommendations were developed.