Arrowtooth Flounder STAR Panel Report

NOAA Western Regional Center
Building 9 Conference Room
7600 Sand Point Way, NE
Seattle, Washington 98115
July 30 – August 3, 2007

Reviewers:

Steve Ralston (panel chair), Scientific and Statistical Committee (SSC) Representative
Patrick Cordue, Center for Independent Experts (CIE, Rapporteur)
Devorah Hart, Northeast Fisheries Science Center (NEFSC)
Jim Ianelli, Alaska Fisheries Science Center (AFSC)
Paul Medley, Center for Independent Experts (CIE)

Advisors:

Brian Culver, Groundfish Management Team (GMT) Representative
Peter Leipzig, Groundfish Advisory Panel (GAP) Representative

Stock Assessment Team:

Isaac Kaplan, Northwest Fisheries Science Center (NWFSC)
Thomas Helser, Northwest Fisheries Science Center (NWFSC)

Overview:

This assessment represents the first dynamic age-structured analysis for arrowtooth flounder. Data sources for the 2007 candidate model included reconstructed catch histories apportioned among three main fleets, the use of 4 survey indices, and length/age data split by sex for selected surveys and fisheries. All indications from the available data and analyses are that there are no conservation concerns. The stock appears to be well above reference points. Except for a number of concerns detailed below, this assessment is acceptable for use by management.

STAT/STAR changes

Throughout the week, the panels refined the assessment which included: 1) changes to the recruitment deviation start year, 2) treatment of natural mortality, 3) refinements to selectivity/retention assumptions about the fillet fishery.

Analyses Requested by the STAR Panel:

Round 1 requests

A: Include a table of landings, including data sources and estimation methods.

Reason: To evaluate sources and magnitude of catches.

Response: To be included in final report.

Discussion/conclusion: Not applicable.

B: How do different discard rate assumptions affect catches in the bycatch fishery. Are results from Sampson's work on historical catch applicable (i.e., Cleaver, flatfish and rockfish catches to get the scale of catches)?

Reason: This is the first attempt to reconstruct the catch time series and all available sources should be investigated.

Response: These were incorporated for subsequent runs.

Discussion/conclusion: The catch time series is still highly uncertain and future analyses should continue to evaluate this uncertainty.

C: Compute an alternative "bycatch" trend based on the ratio of estimated arrowtooth flounder relative to the flatfish complex:

$$B_y = 0.13 \frac{A_y}{S_y} \frac{\bar{S}_{2001-2006}}{\bar{A}_{2001-2006}}$$

where A_y and S_y are the relative biomass levels of arrowtooth flounder and the flatfish species complex (Dover, Petrale, and English soles) and $\overline{A}_{2001-2006}$ and $\overline{S}_{2001-2006}$ are the mean biomass levels for 2001-2006 for those species/species groups. It may be necessary to run the model several times or, preferably, to use smoothed survey estimates.

Reason: To evaluate the static ratio used in the original analysis and to potentially provide a better approach.

Response: The preliminary runs with this evaluation showed that the constant ratio assumption (arrowtooth flounder to other flatfish) is unlikely to hold.

Discussion/conclusion: This approach should be used unless other more quantitative estimates of historical discard levels become available. However, for current management purposes, this issue appears to be relatively minor.

D: Provide Tables and/or Figures of mean lengths over time by sex and surveys.

Reason: To more easily evaluate changes in mean size of arrowtooth flounder caught in the different surveys.

Response: Figures were provided.

Discussion/conclusion: This provided basic background material needed to evaluate selectivity patterns estimated by the model.

E: Evaluate anomalies from triennial surveys for multiple species.

Reason: This issue spans all species for which the triennial survey is used. The change in the timing of the survey to earlier in the year after 1992 could affect the availability of some species to the survey gear.

Response: Preliminary indications suggest that this is an important issue.

Discussion/conclusion: For arrowtooth flounder, the residual pattern of the fit to the triennial survey was reasonably good and the present application in this assessment was seen as acceptable. The influence of the triennial survey as an index was apparently minor. However, further investigation into this issue is recommended.

F: Evaluate IPHC data for arrowtooth flounder abundance indices, if possible.

Reason: This is a scientific survey that operates well within the range of arrowtooth flounder habitat and should be considered as a potential index in future assessments.

Response: This was pursued.

Discussion/conclusion: The IPHC together with the STAT will investigate the potential utility of this survey for future applications. It may be that the period for which species specific information on arrowtooth flounder from this survey is limited.

G: Run the model with catchability (***q***) set to the analytical, unbiased, median method and turn off all priors.

Reason: The effect of priors should be eliminated and estimation of parameters that can be solved analytically should be avoided.

Response: This was done for subsequent models developed during the week.

Discussion/conclusion: The fact that slightly different answers were obtained in this case indicates that there may have been (and there may still be) a problem with convergence.

H: Do a run with sex-specific selectivities where appropriate data are available (e.g., the fillet fishery and the survey).

Reason: There are residual patterns and sex ratio issues that appear to be a cause of some poor quality residual patterns.

Response: These runs were completed and this aspect was retained for the base model recommendation.

Discussion/conclusion: There remained some problems with the fillet fishery selectivity and the ratio of male:female natural mortality that are addressed in subsequent requests.

I: Evaluate a change in fishery selectivity in 2001.

Reason: This seemed appropriate, based on patterns observed in the length frequency data.

Response: The analysis was completed.

Discussion/conclusion: Little effect was noted.

J: Do a profile over male *M*, but with female *M* estimated.

Reason: Due to the constrained selectivity (non-split sex) the original profile was problematic.

Response: This appeared to work appropriately.

Discussion/conclusion: Subsequent requests included this profile with split-sex selectivity options.

K: Do a profile over the length at 50% selectivity for this survey to evaluate this parameter.

Reason: The FRAM slope/shelf survey selectivity is shifted too far to the right (since small arrowtooth flounder are found in shallow regions).

Response: An evaluation with the split-sex selectivity version of the model appeared to rectify this inconsistency.

Discussion/conclusion: This provided further support for the split-sex option on selectivity for the base model configuration.

L: Make a figure of fitted sex ratios by gear type (survey and fishery) obtained from model output and compare to input data.

Reason: Examinations of the length frequency data fits to the model were a cumbersome way to evaluate model predicted sex ratio compared to the observed.

Response: These figures were created.

Discussion/conclusion: The observed variability was quite high compared to model predictions, indicating sampling error was likely quite high. If possible, this type of plot should be included with length frequency figures when sex-specific options are evaluated.

M: Estimate recruitment deviations for the period where data are informative (the alternative method).

Reason: To reduce the number of parameters appropriate to time periods for which data are available to inform the estimates.

Response: The model was very sensitive to this specification.

Discussion/conclusion: As re-requested below, evaluations of 1970 gave very different levels of depletion compared to models with all recruitment deviations were estimated. Further investigations were encouraged with a model that was closer to a final "base" model.

N: Do a run with no fishing, with parameter values fixed from runs with proposed catch histories.

Reason: Since catch history is highly uncertain, the relative impact of different assumptions about catch history can be reasonably evaluated.

Response: This is an output option within SS2.

Discussion/conclusion: This is a second order issue but may help in future evaluations of

uncertain catch histories.

Round 2 requests

O: Create a new baseline with the following core specifications: all priors removed, analytical catchabilities, added catch data recommended from B. Culver from Cleaver, and with recruitment deviations starting in 1970. Convergence diagnostics for this model should be evaluated, including bounds checking, checks for derivatives and the maximum gradient.

Reason: After discussions, these features seem most appropriate.

Response: This was done and results guided subsequent discussions.

Discussion/conclusion: see below

P: As in O:, but with split-sex selectivity.

Reason: After discussions, this configuration seems most appropriate.

Response: As in H above,

Discussion/conclusion: See under H above.

Q: Start the vector of recruitment deviations earlier than 1970 for comparisons.

Reason: After discussions, this sensitivity analysis was deemed to be useful.

Response: As in M above,

Discussion/conclusion: See under M above.

R: Do a GLM with seasonal fixed effects for triennial surveys

Reason: There is a concern that the shift in timing for this survey, particularly between the early and later periods, may affect the availability of arrowtooth flounder to the survey.

Response: A GLM was run where average Julian day anomalies (relative to overall mean) were computed and applied as an independent continuous covariate. Estimates of annual coefficients with and without the date covariate were inconclusive.

Discussion/conclusion: The seasonal effect could potentially be tested with the FRAM survey data which occurs over more months in multiple passes (N-S). This was considered a general future research recommendation. For the purposes of the arrowtooth flounder assessment and management outlook, the current treatment (inclusion of the triennial index) is appropriate.

S: Provide year-specific discard ratios for this and other studies (e.g., Pikitch etc).

Reason: As in earlier requests, there are concerns over the catch time series, particularly regarding estimated discard levels.

Response: Provided in a table for panel review.

Discussion/conclusion: The levels of bycatch seemed to be within the ranges evaluated (e.g., under request A).

Round 3 requests

T: Do a full integrated GLMM with seasonal fixed effects for triennial surveys and compare with and without the seasonal effect.

Reason: There is a concern that the shift in timing for this survey, particularly between the early and later periods, may affect the availability of arrowtooth flounder to the survey.

Response: The request was not completed due to difficulty in obtaining the data in an expeditious manner.

Discussion/conclusion: The survey data should be made readily accessible to all stock assessment authors so that timely analyses can be conducted.

U: For a candidate "new" base model, use sex-specific selectivity and remove the retention curve from the fillet fishery. Use the landings and the discard ratios to generate total catch estimates for this fishery. Recruitment deviations are to start in 1970.

Reason: There is a component of discarding that isn't size based and (perhaps) this is causing the selectivity estimates to be unreasonably shifted to larger arrowtooth flounder.

Response: More time will be needed to do this properly (see R: above).

Discussion/conclusion: This is potentially a general issue for many groundfish stocks.

V: For the "new" candidate base model (sex-specific selectivity), evaluate the effect of starting the recruitment deviations vector in 1960, both with and without the "zero sum" option invoked.

Reason: To decide if 1970 is the most appropriate start year for recruitment deviations

Response: Based on the evaluation, the zero sum option had very little effect.

Discussion/conclusion: 1965 appears to be an appropriate compromise between estimating

recruitment deviations in each year (dating back to 1916) and a more recent year (e.g., 1970).

W: Conduct profiles over alternative *M* values for one sex, allowing the other to be freely estimated. For both sexes, keep *M* constant with age (no difference between young and old).

Reason: There was little biological rationale provided to suggest that ages 0-4 yrs should have a different *M* than the older ages.

Response: This was done for all subsequent evaluations of "base" models

Discussion/conclusion: The stability of the model apparently changed with this modification. There is some concern that convergence problems may be persisting.

X: Based on sensitivities and profiles, propose a base model.

Reason: To advance the stock assessment for management purposes

Response: The needed evaluations were completed.

Discussion/conclusion: see AB below.

Y: Adjust bycatch ratios based on the new base model. Iterate to see if it changes. Use all available information for discard rates.

Reason: See C above.

Response: See C above.

Discussion/conclusion: See C above.

Z: Do a sensitivity run with the Triennial survey catchability split in 1994 to account for the possible effect of a shift in survey timing.

Reason: The timing of the survey changed and could impact trend information for this survey.

Response: Time was unavailable to complete this during the week of the review.

Discussion/conclusion: This appears to be a relatively minor issue for arrowtooth flounder, but should be considered in future assessments/evaluations. See AC below.

AA: Explore axes of uncertainty, in particular, a profile over alternative values of R_0, perhaps with *M* freely estimated.

Reason: To provide an evaluation of principal axes of uncertainty.

Response: Further work on a base model is needed before this can be appropriately obtained (uncertainty evaluation).

Discussion/conclusion: See AB-AD below.

Round 4 requests

AB: Given the apparent status of the stock and the lack of conservation issues with this stock, a base model recommendation is to retain split-sex selectivity, ignore information on retention rates, start the recruitment deviations in 1965, fix peak selectivity of the fillet fleet to 60cm and estimate natural mortality.

Reason: This appears to be a suitable base model specification for management purposes.

Response: Fixing male natural mortality at 0.274 stabilized the estimation properties and results appear to be reasonable.

Discussion/conclusion: See U above regarding the issues related to retention-rates.

AC: Same as AB, evaluate model sensitivity to the triennial survey index of biomass.

Reason: Since analyses on splitting this survey into two periods of availability were impractical, this will provide some indication on the importance of the survey to this assessment.

Response: There appeared to be very little impact of this survey on key model results.

Discussion/conclusion: This indicates that the constraints on natural mortality combined with catch history and stock-recruitment steepness affect model results most significantly.

AD: Produce alternative scenarios that capture the uncertainty as follows: bracketing uncertainty in *M* combined with low and high catch histories to pick extreme cases. The bracketing of uncertainty in *M* should capture approximate lower and upper 25^{th} percentiles of probability estimates.

Reason: This appears to be the main axis of uncertainty.

Response: To be determined.

Discussion/conclusion: Not applicable.

Description of final base model:

The final base model included all data from the pre-STAR base model, with the addition of the

Cleaver (1951) 1928-1949 Oregon catch data for the mink food fleet. The full catch history for the three fleets covered 1928-2006. Commercial length data were available for 1986-2006, with age data for 1986-1991, 1998, and 2001-2005. The model included length data and abundance indices from the NWFSC Slope-Shelf Survey (2003-2006), the Triennial Shelf Survey (1980-2004), and the AFSC Slope Survey (1997, 1999-2001). Survey indices of abundance were also incorporated from the NWFSC Slope Survey (1998-2002), as were ages from the NWFSC Shelf Slope Survey (2003-2006).

The final model specifications included:

- catchabilities (***q***) calculated analytically as median unbiased;
- steepness of 0.902 (from Dorn's meta-analysis);
- recruitment deviations estimated from 1965 on;
- split-sex selectivity;
- full retention for the fillet fleet (i.e. discards incorporated into catches based on observed discard rates, rather than estimating discards within the model);
- selectivity peak parameter for the fillet fleet fixed at 60cm;
- natural mortality fixed at 0.166 for females and 0.274 for males (based on Hoenig 1983 for females and likelihood profiles for males);
- sigmaR=0.8 (tuned);
- tune CVs and effective sample sizes;
- no priors.

Comments on the Technical Merits and/or Deficiencies of the Assessment:

Technical Merits:

- The preparation of documentation for the panel was very good, particularly since this was the first age-structured model developed for arrowtooth flounder.
- The assessment was based on SS2 software, which has been accepted as the standard for west coast groundfish.
- The method of including age observations as conditional on the length was considered a better way to include age data in the model, thereby avoiding *ad hoc* weighting schemes.
- The STAT team was very responsive to the STAR panel's requests.

Technical Deficiencies

- The catch time series is highly uncertain.
- The treatment of the triennial survey should account for the change in the timing of the survey between the early and later parts of that time series.
- Some selectivity parameters appeared to be inestimable.
- There appeared to be some difficulty in obtaining proper convergence of the model.

Explanation of areas of disagreement regarding STAR Panel recommendations:

Areas of disagreement among the members of the STAR Panel:
There were no areas of disagreement among the five panelists.

Areas of disagreement between the STAR Panel and the STAT team:
There were no areas of disagreement between the STAR panel and the STAT team.

Unresolved Problems and Major Areas of Uncertainty:

This section focuses on major uncertainties associated with the arrowtooth flounder assessment.

- Catch histories for arrowtooth flounder are problematic and require further study.
- Recruitment deviation start time—the model was very sensitive to this specification.
- Natural mortality rate is poorly known.
- Selectivity and retention in the fillet fishery is poorly specified.
- The triennial survey appears to have little impact on this assessment. However, it appeared that in some cases (e.g., when natural mortality was estimated), the estimates became unstable (i.e., an unreasonably large stock size resulted). This is cause for concern.

Concerns Raised by GMT and GAP Representatives During the Meeting:

The GAP and GMT representatives did not raise any objections to the discussion and outcome of the arrowtooth flounder STAR panel review, but wished to make note of the following points:

- There is a general need to conduct a comprehensive examination of historical catch data.
- Access to NWFSC trawl survey data is unnecessarily restrictive.
- Trawl logbook data should be examined and utilized to a greater extent.

Recommendations for Future Research and Data Collection:

For the next arrowtooth flounder stock assessment

- The arrowtooth flounder catch history should be reconstructed using all available data including catch by gear and by region. The reconstruction should include an envelope of high and low values to set bounds for exploration of alternative catch histories. As has been recommended previously by a variety of STAR Panels, the reconstruction of historical landings needs to be done comprehensively (i.e., with other species) to ensure efficiency and consistency.
- Evaluate the feasibility of a bi-lateral assessment with Canadian scientists, perhaps through the TSC (Technical Subcommittee of US Canada groundfish working group).
- Investigate the importance of calendar date on catch rates from the triennial survey and propose an adjustment, if needed.

Generic issues for groundfish assessments

- Establish a *meta* database of all data relevant to rockfish stock assessments. The database should include enough detail about the nature and quality of the data that a stock assessment author can make a well informed decision on whether it could be useful for their stock assessment.
- Establish *accessible* online databases for all data relevant to groundfish stock assessments, so that assessment authors can expeditiously obtain the *raw* data if required.
- Establish a database for historical groundfish catch histories, “best” guesses and

estimates of uncertainty (and processes for updating and revising the database).

- Develop a concise set of documents that provide details of common data sources and methods used for analyzing the data to derive assessment model inputs.
- Develop standard and appropriate methods for modeling age and length data, including choice of distribution, initial variance assumptions, and tuning methods (current methods can and should be improved).
- Routinely produce and present supporting documentation for any derived indices which are included in a stock assessment model (e.g., GLMM derived trawl survey abundance indices).

Acknowledgments:

The STAR panel would like to compliment the NWFSC, especially Stacey Miller, for coordinating the meeting and the review of the arrowtooth flounder stock assessment. The STAT team (Isaac Kaplan and Tom Helser) is also commended and thanked for undertaking a wide variety of analyses, all on short notice, which insured the successful completion of the review.