Dover Sole

STAR Panel Meeting Report
June 20-24, 2005

NOAA Fisheries
Northwest Fisheries Science Center
Hatfield Marine Science Center
Newport, Oregon

STAR Panel:
Tom Barnes (Chair) – SSC representative
Selina Heppell – Oregon State University
Bob Mohn– Center of Independent Experts (outside reviewer)
Stephen Smith (Rapporteur) – Center of Independent Experts (outside reviewer)
Grant Thompson – NOAA Fisheries, AFSC

PFMC:
John Field – GMT representative
Rod Moore – GAP representative
Mark Saelens – GMT representative

STAT:
David Sampson – Oregon State University

Overview

The STAR Panel reviewed an assessment of the status of the Dover sole *(Microstomous pacificus*) stock off the U.S. Pacific coast. The review was conducted at the Hatfield Marine Science Center in Newport, Oregon during the week of June 20 to 24, 2005. This stock of Dover sole was last assessed in 2001.

This assessment used the new version of the Stock Synthesis program (SS2 version 1.19). The length and age composition data were separated into two fisheries: a northern fishery operating in the US Vancouver and Columbia INPFC regions and a southern fishery in the Eureka, Monterey and Conception regions. The period modeled in the assessment extended from 1910 to 2004 with fishing beginning in 1911. Data in the assessment model included fishery length composition data from 1966 to 2004 and fishery age composition data from 1981 to 2004. The biomass indices were derived from trawl logbook catch rates (1978 to 1995), and biomass estimates and length and age composition data from bottom trawl research surveys of the shelf (1980 to 2004) and slope (1992 to 2004). Retention and discarding were modeled using logistic functions of length.

The presentation of the stock assessment began with a discussion of the major unsolved problems and issues that were faced in preparing this work. Some of the issues were similar to those raised in the previous assessment in 2001 and concerned the fact that the size and sex distributions of Dover sole are highly variable by depth and between INPFC areas and have changed over time. These patterns and other size related issues, such as a tendency for the fish discarded in the south fishery being larger than those retained were difficult or impossible to model in the stock synthesis program.

An investigation of the NWFSC survey data for all species reviewed at this STAR panel indicated that the use of the mean density in the area north of Point Conception to estimate densities in the south for years that the survey did not go below the point was not supported by recent data. Biomass indices from this survey series were recalculated with the Conception stratum being confined to the area north of Point Conception.

A great of deal of time was spent by the STAT member and the panel trying to develop a base model but convergence issues complicated arriving at a decision. In the end, a slightly modified version of the preliminary base model originally proposed in the assessment was accepted as the best that could be developed with the data and model at hand.

The stock has exhibited an increasing trend since about 1995, with the depletion for 2005 estimated to be 0.63.

Analyses requested by the STAR Panel

1. Recruitment deviations were included for the period 1910-2003 in the preliminary base model but there seemed to be little variation in recruitment until sometime between 1930 and 1940. The panel requested that the starting date for recruitment

deviations be changed to this later period with the actual date to be judged by plotting recruitment deviation standard errors against year.

The plot showed that the estimated standard errors were equal to the assumed sigma (R) (0.35) for the first 20 years of the modeled period and did not begin varying from this value until about 1930. Based upon visual inspection of the plot of recruitment deviation standard errors time series, it was decided that 1930 would be an appropriate year for starting the recruitment deviations.

2. Estimates of length-at-peak for fishery selection curves were estimated as being very close to 36 cm for both North and South fisheries. The author proposed fixing the length-at-peak to 36 cm for both fisheries to simplify the model. The panel requested that the author try a profile of a range of sizes (34 to 40 by 1 cm) to assess the impact of fixing length at 36 cm.

   The profile was conducted over the range from 34 to 40, as requested, but by increments of 2 because that is the bin width of the length composition data. The model failed to converge at 34 and 40 cm. The best fit to the data occurred when the value of the female length-at-the-peak parameter was 36; the likelihood for this run was only slightly different from the value obtained with the preliminary base run. There was little evidence overall to suggest that setting length-at-peak for both areas at 36 cm would be a problem.

3. The original presentation of the assessment had included evaluating a profile on steepness parameter ranging from 0.7 to 0.95. This analysis showed that there was little change in log likelihood over this range of steepness parameter. The panel requested that the profile be extended to investigate the impact of setting the steepness parameter to a lower bound of 0.4. This analysis should be done using version of the model with recruitment deviations starting in 1910 for comparison purposes.

   Decreasing the steepness parameter to 0.4 resulted in a total log likelihood of 1370.8 compared to 1362.4 for steepness of 0.8., which indicates that the data do not provide much information about the value of the steepness parameter. What little information there is, however, suggests that steepness is a relatively high value. The estimate of percent exploited decreased from 61.3 to 49.8% for the lower steepness level. The panel discussed including the uncertainty about steepness in the decision table for this stock.

4. Selection deviations were included annually starting in 1970 to account for time-variation in selection. However, there appeared to be three distinct time periods (early, mid-years and post-1995) in the selection variation. The panel requested that the selection deviations be included in three time blocks to simplify the model and better reflect the patterns in the data.

The three time periods were defined as pre-1980s for the early period, the middle period of 1980 to 1995 and late corresponded to 1995 to present. The model with three time-blocks had only 121 estimated parameters (versus 209 in the preliminary base model) and produced a total negative log-likelihood value of 1492.0 (versus 1313.5 for the preliminary base model). Despite this change, the panel agreed that having selection deviations confined to time periods was more realistic and that the simpler model with three time-blocks should be preferred.

5. A diagnostic plot of the time series for the four survey indices of biomass along with the model estimate of biomass was requested, where each index has been adjusted for its corresponding catchability coefficient (q). This was done to evaluate how the survey data was influencing the final model estimate.

   The biomass trend and survey trends were very similar with the increase in biomass in the recent years reflecting trends in the AFSC and shelf survey.

6. The version of the model with recruitment deviations starting in 1930 and selection deviations in three time period blocks seems to be a potential base run. Profiling this version against a range of steepness parameters (h) indicates that this base run should also include h set to 0.8. However, there were concerns about 0.8 being a local minimum and the panel suggested that the profiling against a range of h from 0.4 to 0.95 with increments of 0.05 might help ascertain if the model solution at h=0.8 was a global or local minimum.

   The profile investigation of steepness indicated tension between the length-at-age information from the AFSC slope survey and the NWFSC slope survey (see 10 below). It also seemed to indicate fairly large sensitivity to the steepness parameter (a change of 40 log-likelihood points when steepness changed from 0.95 to 0.425). However, the log-likelihood surface appeared to vary erratically, which suggested that some of the runs might have converged to local minima. The panel suggested that the profile investigation be expanded to include investigation this issue.

7. The panel suggested model runs with randomization of initial values for the potential base run discussed above as a way of addressing local minima concerns.

   Forty model runs were made with perturbed initial parameter values. One run did not converge (large maximum gradient), 15 runs converged to the same value log-likelihood value as the original Run 219, and 20 runs converged to negative log-likelihood values that were almost 5 units better than the Run 219 value. These findings seemed to indicate some problems here, but discussion about changing the survey series subsumed this request for now (see 11 below).

8. In the original presentation, the author had to use different values of lambda for the different likelihood components to get convergence. The potential base model has changed enough and the panel suggested that it was time to re-visit the question of setting the lambdas to be different than 1.0 for all components. The

panel requested that a run of the potential base model be done with all lambdas equal to 1.0.

The newer model would not converge when all the lambdas were set to 1.0 (large maximum gradient).

9. Concern was raised by the panel about the NWFSC survey biomass estimates for the Conception stratum. Prior to 2002, there was no sampling below Point Conception and the mean density north of Point Conception was used for the area south of the point. An investigation of density estimates above and below the point for 2002 to 2004 when there were stations in the south indicated that the density in the south was much smaller than in the north. Tom Helser developed new biomass indices using only the area and stations north of Point Conception. The panel directed the author to refit the model to this new survey series realizing that there would not be time to adjust the survey length composition for the change in the survey series.

   The results from the model run (recruitment deviations starting in 1930 and selection deviations in three time blocks) with the new data were not very much different than those with the previous NWFSC survey data with respect to biomass, depletion, etc.

10. The profile investigation of steepness in 6 above indicated that there seemed to be tension between the length-at-age information from the AFSC slope survey and the NWFSC slope survey. The panel asked the author to investigate the effect of changing lambda for the AFSC survey information from 0.10 to 0.15 for new data.

    This investigation was suggested to resolve problems about finding a minimum when profiling the model for different values of steepness. This work was not done but insight about using 0.15 instead of 0.10 was obtained in the analysis conducted in 11 below.

11. The panel suggested model runs with randomization of initial values for the potential base run with the new survey data as a way of addressing local minima concerns. These runs should be done for two choices of lambda for the AFSC length-at-age data discussed above.

    Randomization of starting values continued to show that the model had difficulty converging. Log likelihood varied from approximately 1497 to 1516 while virgin spawning stock biomass changed by around 3 percent. Adjusting the lambda for the AFSC slope survey length-at-age information from 0.10 to 0.15 did not seem to help this situation.

12. A closer look at the results from randomizing the starting values indicated that most of the change was associated with the fishery and survey length compositions. A run with these compositions down-weighted was recommended to see if the situation was improved.

    The results from this were difficult to interpret because recruitment deviations were reduced considerably and the fit to the commercial CPUE data became stronger.

13. The panel decided to continue investigating the sensitivity the model with recruitment deviations starting in 1930, selection deviations in three time blocks and the new survey data and requested runs profiling on M with h ranging from 0.4 to 0.9 and incremented by 0.1.

    The profile for steepness did not help resolve a solution for steepness and given the problems with convergence indicated earlier by the randomization runs, the panel decided to stick with h=0.8. Overall, there was little variation in spawning stock biomass and depletion with the different values of steepness.

    The profile for natural mortality was not very informative either although there was more variation in population estimates with different M than observed for the profile for h. The panel decided to accept the model with M=0.09 as the base but use M equal to 0.07 and 0.11, as representing our uncertainty with respect to low and high productivity, respectively. Forecasts at 40:10 were made with the three states of nature defined by the values of M above.

Final base model description

The final Base Model was defined with h=0.8, M=0.09 and catchabilities (q's) estimated. The different weights (lambdas) for the likelihood components were kept as the author had set them.

As a result of model and data explorations requested by the STAR panel, the final Base Model configuration and data were modified from that presented in the Review Draft assessment document as follows:

- Recruitment deviations were permitted for the period 1930-2003.
- For the fishery selection curves the parameters for the female length at the peak, which were estimated in the preliminary base model, were fixed at 36 cm for north and south fisheries.
- In both fisheries the parameter controlling the ascending inflection point, which had annual deviations in the preliminary base model, was allowed to vary between three time periods (1910-79, 1980-95, 1996-2004) but was constant within each time-block.
- The final base model used the revised NWFSC slope survey biomass estimates (provided by Tom Helser during the STAR week).

- The final model run was the one with the best fit selected from a set of 40 runs that were generated with random "jitter" on the initial parameter values.
- The runs for the alternate states of nature were the runs with the best fit selected from sets of 40 runs that were generated with random "jitter" on the initial parameter values.

Following the adjournment of the STAR panel, a decision table was produced in which the states of nature were based on a range of natural mortality of 0.07– 0.11. The alternative management decisions consisted of the catch time series corresponding to low, medium and high catch scenarios. The low catch was set equal to the average of 2000-2004, the medium catch was two times the low catch, and the high catch was F40% yield using the 40:10 rule for the Base Model (M=0.09).

Technical merits and/or deficiencies in the assessment

The panel appreciated David Sampson's patience and dedication in responding to the many requests for further analysis. Many problems with this assessment remain and are discussed in last two sections of this report.

Areas of disagreement regarding STAR Panel recommendations

There were no significant areas of disagreement.

Unresolved problems and major uncertainties

The size and sex distributions of Dover sole are highly variable by depth and between INPFC areas and have changed over time. It is difficult to determine whether these variations are due to differences in size-related discarding or to differences in selection, related either to gear or to depth of fishing. The size-discards and size-selection effects are confounded in the fishery size composition data. Only a few observations are available for the size-distributions of discarded fish.

The Observer Program data indicated latitudinal differences in the pattern of discarding of Dover sole caught in deep water (> 300 fm) with the discarded fish in the south being slightly heavier on average than the retained fish, while the discarded fish in the north (US Vancouver and Columbia) are lighter than the retained fish.

The current version of Synthesis cannot accommodate size-selection curves in which peak selection occurs at different sizes for females versus males as was evident in the length composition data from the trawl surveys and the fisheries.

The assessment was unsuccessful in configuring a model that resolved the conflicting signals evident in the length composition data versus the age composition data versus the length-at-age data.

The weights associated with the likelihood components (lambdas) needed to be set to different values to get the preliminary base model to converge. The panel was uncertain about the impact of doing this on the reliability of the estimates from the model.

Recommendations for future research

1. Investigate model structure to diagnose and solve convergence problems. Some of the following recommendations are considered elements of this investigation.

2. Develop a model for this population to overcome limitations with handling size/sex related patterns either by having SS2 modified or use some other approach.

3. Collect more information on length composition (by sex if possible) of discards, especially in the southern area.

4. Determine factors underlying discard patterns in north and south fishery. Factors such as change in acceptable size to markets, targeting by depth, problems with jellied condition of the larger fish in the south and changes in regulations (e.g., varying trip limits) were all suggested as having an influence.

5. Explore having the CV of length-at-age interpolated being a function of age instead of length.

6. The commercial CPUE is only used up to 1995 because of problems with changes in regulations after this time. Extension of this series should be investigated by determining how this index could be used as a biomass index accounting for problems with trip limits, bycatch limits, etc.