Stock Assessment Review (STAR) Panel for
Dover sole, Sablefish, Longspine thornyhead, and Shortspine thornyhead

June 20-24, 2005
Hatfield Marine Science Center
Barry Fisher Building, Main Conference Room
Newport, Oregon

6 July, 2005

Stephen J. Smith
91 Castlewood Drive
Dartmouth, Nova Scotia
Canada, B2V 2R4
902-435-7843 (residence)
902-426-3317 (office)
smithsj@mar.dfo-mpo.gc.ca

Executive Summary

Stock status for Dover sole, Sablefish, Longspine thornyhead, and Shortspine thornyhead in 2005 were reviewed by a STAR Panel. These four species are caught in mixed species fisheries and trip limits for any one species could have an effect on the fishery for the other species. All of the assessments incorporated a new survey series for the slope developed with a new method which allows for a common abundance trend to be estimated from multi-vessel surveys. These stock assessments represented the best that could be done in the time available and with the data at hand. However, determination of stock status for each of the species came down to subjective decisions about one or more values of natural mortality rate, the steepness parameter of the recruitment model and the catchability of the fish to the survey gear. More basic research on the population dynamics is needed for the species to reduce this subjectivity in the future.

Background

A STAR Panel was convened in Newport Oregon form June 20 to 24 to review the stock assessments for Dover sole, Sablefish, Longspine thornyhead, and Shortspine thornyhead in 2005. Guidelines for the STAR Panel stipulate that a maximum of two stocks will be considered by a panel but these four species (DTS) are considered together because they are usually caught together and trip limits for anyone species could constrain the catch of the other species. Participants are listed below. Members of the STAR Panel were in attendance the full week and STAT members were present for much of time dependent upon the amount of extra work they had been assigned. Attendance by the other participants depended upon their own schedules and interest.

Participants:

Star Panel:

Tom Barnes, Scientific and Statistical Committee (SSC) Representative, STAR Panel Chair
Grant Thompson, Alaska Fisheries Science Center (AFSC, Rapporteur: Sablefish)
Selina Heppell, Oregon State University (OSU)
Robert Mohn, Center for Independent Experts (CIE, Rapporteur: Shortspine thornyhead)
Stephen Smith, Center for Independent Experts (CIE, Rapporteur: Dover sole)

Rod Moore, Groundfish Advisory Subpanel (GAP) Representative
Mark Saelens, Groundfish Management Team (GMT) Representative
John Field, Groundfish Management Team (GMT) Representative (Rapporteur: Longspine thornyhead)

Stock Assessment Teams members present:

Dover sole - David Sampson, Oregon State University (OSU),
Sablefish - Michael Schirripa and J.J. Colbert (OSU), Northwest Fisheries Science Center
Shortspine thornyhead - Owen Hamel, Northwest Fisheries Science Center
Longspine thornyhead - Gavin Fay, University of Washington

Pacific States Marine Fisheries Commission:
Susan Coccetti
Omar Rodriguez

Fishing Industry:

Coos Bay Trawlers' Assoc/Bandon Submarine Cable Council:
Steve Bodner
Gerald Gunnari
Brad Pettinger

Fishermen's Marketing Association:
Peter Leipzig

Makah Tribe:
Brandon Bryant
Steve Joiner

Oregon State University:
Vladlena Gertseva
Maria Jose Juan Jorda

Description of review activities

Most of the materials required for the STAR Panel were distributed either by email or on a CD sent via FEDEX on June 8, 2005. The CD included the draft stock assessments for

the four species and a very comprehensive set of background information including documents on survey estimates, catch rate analyses, previous assessments, and findings from the data and modeling workshops (see Bibliography). The CD also contained the executable file, manual and example files for Stock Synthesis 2 the software package used for all four species. While I did not get a chance to run the program, I found the manual and the example files to be very helpful in understanding how the assessments were conducted. Stacey Miller (NOAA Fisheries, NWFSC, Seattle) had emailed Helser et al.'s (2005) document on the survey estimates on June 8 because I was away at-sea from June 9 to June 11 and would not be around to receive the CD when it arrived. I wanted to pay special attention to Helser et al.'s (2005) derivation of the new survey estimates that would be used for all four assessments.

The chair of the panel, Tom Barnes, assigned rapporteurs when he sent out the agenda on June 8. I was assigned to Dover Sole and concentrated my attention on this stock assessment while reviewing those for the other three species. The Dover Sole assessment on the CD was incomplete and was missing a number of tables and figures including those summarizing the discard information from the observer data, age reading error and coefficients of variation in the length-at-age data. A number of the tables that were included were not up-to-date (e.g., Table 1 Management regulations). There was information on the elements included in the proposed base model but there was nothing on the results, including sensitivity analyses, reference points or decision tables.

An addendum for the Dover Sole assessment was sent out by email on June 13 that contained most of the missing material from the CD version of the stock assessment. However, forecasts were not included. Paper copies of supplemental material were handed out to the panel on the first day of the meeting (June 20), containing a large number of sensitivity and retrospective analyses plus a comparison with the previous assessment.

The CD versions of the stock assessments for the other three species were complete. Bob Mohn sent a series of analysis suggestions for Shortspine Thornyhead to the Chair on June 15 and received a detail reply from the author, Owen Hamel the same day. This correspondence was also distributed to the panel that day. On June 17, we were also sent an exchange between Michael Schirripa and Martin Dorn on the Sablefish assessment. All of this information helped me get up-to-speed on these assessments.

We started off with the Sablefish assessment on the morning of the first day of the meeting but broke off from the presentation at 9:30 h to conduct a video conference with the office in Seattle in which Tom Helser gave a presentation on the model-based estimates of abundance for the NMFS (AFSC and NWFSC) slope stratified random surveys used for all four assessments. There are three main elements to this approach:

1. Following the work of Pennington (1983) and others, model the survey data as a two-stage event of catch and no catch for each species with the numbers or biomass of animals caught in the first stage modeled using either a Lognormal or an exponential family distribution, in this case Gamma, or inverse-Gaussian. The

two-stage aspect of catch/zero catch was modeled using a Bernoulli distribution. An empirical evaluation of the observed distribution of the non-zero catches suggested that the Gamma distribution fit the best, similar to findings for haddock by Stefánsson (1996).

2. Commercial fishing vessels have been used for the NWFSC surveys with different vessels being used over time. The problem of obtaining a survey series independent of or corrected for vessel changes over time was solved by characterizing vessels as a random effect within a generalized mixed-effects model (GLMM) approach.
3. For the simple case of one vessel and no covariates, the estimates for the two-stage event model using the lognormal distribution, the so-called Δ-distribution are given by Pennington (1983). However, analytical solutions for the estimates from the Gamma model with the random effects are not available and the authors used a Monte-Carlo Markov Chain approach to solve for the product of the Bernoulli distribution and the GLMM with the Gamma error distribution (Helser et al. 2004).

In my opinion, the latter aspect of the method is as important as all the others but very little detail about it was provided in the published paper (Helser et al. 2004) or the document sent on the CD (Helser et al. 2005). Fortunately, Tom Helser was able to provide explanation and details during the video conference and afterwards to email queries that I sent him which were in turn provided to the panel.

The Dover Sole assessment was presented on Monday afternoon and the two thornyhead species were covered on Tuesday. The one thing in common amongst all of the assessments was that the results from the stock Synthesis model were extremely sensitive to assumptions about either natural mortality or the steepness parameter of the recruitment curve or the catchability coefficient for the slope survey or some combination thereof. Much of the extra work assigned to the STAT team members was to try different model formulations to determine the best values for these parameters.

During the initial presentations there was some discussion about how the slope survey estimates were adjusted for the recent extension of the NWFSC survey to the area south of Point Conception, California. Prior to 2002, coverage of the Conception Stratum off of California in the NWFSC survey series had been limited to the area north of Point Conception. Biomass estimates for the whole Conception stratum for surveys prior to 2002 were calculated by using the density from area north of Point Conception for the whole stratum area. This approach assumes that densities are similar in the north and south areas of this stratum. This assumption is particularly important given that more than 70 percent of the stratum area is south of Point Conception and this area is large relative to the other strata in the survey.

A comparison of the longspine thornyhead densities north and south of Point Conception from surveys for 2002 to 2004 when both areas were sampled indicated that densities in northern area ranged from 0.15 to over 10 times the density in the southern area for that time period. The panel concluded that it was inappropriate to assume that the densities

north of Point Conception could be used to represent the whole Conception stratum for the surveys prior to 2004. Tom Helser recalculated the survey indices using the Delta-GLMM approach limiting the Conception stratum and the survey tows to be the area north of Point Conception and provided these new data to the STAT team members on Wednesday. All of the assessments except for shortspine thornyhead were revised for these new indices. The survey length compositions should also be corrected for these changes in stratum definition but there wasn't time to make this change during the meeting. In the case of shortspine, survey densities north and south of Point Conception for 2002 to 2004 were very similar and therefore no changes to the survey series were made.

We spent the remainder of the week reviewing alternative model runs, evaluating the impact of the new survey series and trying to develop a base run for each species. The base runs were all developed by Friday at noon and the panel spent the afternoon editing the individual STAR Panel reports for each species. Tom Barnes has taken these reports for final editing and plans on distributing the penultimate versions by July 8, 2005.

Summary of findings

Primary sources of uncertainty in the assessment

For each assessment, determination of the base model came down to making assumptions about probable values for either natural mortality (M), steepness (h) or catchability (q) or some combination thereof. In each case, parameters were fixed or had a restrictive prior distribution set on likely parameter values. Setting a value or restricting the range of likely values for h can make any estimates of maximum sustained yield (MSY) targets and associated quantities arbitrary or unreliable, necessitating the use of proxies (e.g., 40% of the unfished spawning stock biomass for the biomass associated with MSY). Estimates of M and q affect the estimate of absolute biomass, used to determine if the stock is overfished.

It was pointed out by one of the STAR Panel members that research recommendations are often directed towards obtaining more data or more model development but do not generally focus on specific research aimed at reducing our uncertainty about M, h or q. For our part we tried to focus on specific recommendations directed to a better understanding the possible ranges for these parameters and more work on selectivity and catchability along the lines of Lauth et al. (2004) was encouraged, especially for the thornyhead species.

While the Stock Synthesis program allowed for different weights for each of the likelihood components only the Dover Sole assessment chose to use different weights. These were somewhat arbitrarily chosen so that the model converged. As far as I know there are no guidelines available on how to optimally choose weights or whether having weights other than unity (1.0) will really result in a maximum likelihood solution.

Strengths and weaknesses of current approaches

All of the stock assessments used Stock Synthesis 2 to model the population dynamics and estimate management parameters. While this version of the software seems to be a major advance over the previous version there were still important features of the population dynamics that could not be accommodated. In particular for Dover sole, the software could not take into account that there were different selectivity curves for males and females and that the mean size of discarded fish in the southern fishery exceeded the mean size of the retained fish. The Dover sole stock assessment was not ready in time and the STAT member reported that he had been unable to configure a model in Stock Synthesis that handle the highly variable size and sex distributions that differed between areas and over time.

One of the strengths of the Stock Synthesis software is that it can easily integrate a number of diverse sources of data and understandably STAT members try to use as much data as they have available. However, it is difficult for reviewers to evaluate whether or not all of the data is really useful in a 5 day, four species panel meeting. In addition, the ability to use so much data and structure does come at the cost of increased computing time being required. This was particularly problematic for the Dover sole assessment where model runs were taking almost an hour. As a result the STAT member for Dover Sole was away from the meeting much of the time overseeing the model runs and missing out on much of the discussion of model and data issues for the other species that may have been relevant to Dover sole as well.

For the most part, the STAT members experience with Stock Synthesis 2 was quite recent and therefore limited and there were issues that we could have used some insight on. For example, some of the assessments included forecasts during their routine runs and we noted that there was a forecast recruitment likelihood component reported in the output. While we did determine that this component was an independent additive component, we did not determine why or how forecasts would have a likelihood component associated with them. We did not find any discussion of this issue in the Stock Synthesis documents on the CD. Another issue concerned the use of priors for some of the parameters. It appears that in the maximum likelihood mode, these priors may work as penalties on the total likelihood function but it was not clear if the resultant parameter estimates would be the same as the posterior means or medians when a full Bayesian approach using MCMC was conducted. This information may be available somewhere but was not available to us.

Finally, I am concerned that even though the STAT members have a number of data sources to use and a sophisticated software package to integrate these data, the public perception may be that the final determination of stock status is quite arbitrary. That is, we spent a fair amount of time tweaking parameters for which there is little information in the data (i.e. M, h, q) or discussing the setting of weights for the likelihood components for Dover sole. More than one member of the fishing public present at the meeting had difficulty trying to line up their perceptions of stock status with our discussions on M, h and q or likelihood weights. This was especially problematic when a

comparison of observed and fitted survey estimates indicated lack of fit as they did for sablefish. I would recommend that every consideration of an alternative model formulation should be accompanied by an evaluation of how well the model fit the data.

Conclusions/recommendations

The stock assessments represented the best that could be done in the time available and with the data at hand. However, it was frustrating that the process came down to subjective decisions about possible values for natural mortality, steepness and catchability because the models could not estimate all or some of these parameters. I propose the following recommendations for future STAR Panel meetings for these or other species:

1. Discussions about possible values for model parameters seem to be disengaged from possible impacts of changes to the model on fits to the data. It would be helpful to address these changes in terms of the goodness of fit for panel members and for members of the public attending the meeting who may not have the background to follow the modeling discussions.

2. Stock Synthesis allows analysts the ability to incorporate data from many diverse sources. The impact of using all of these data sources could be better assessed if the results of fitting some basic models (e.g., production models) with a minimum set of trend data (e.g., surveys, commercial catch rate) could be presented for comparison. For example, if the main difference between the basic model and the final assessment model was that the latter included area specific size composition data and as a result the trends of the two models diverged substantially, attention could be paid to the reasons for the change and the assumptions underlying the use of the size composition data.

3. Computing facilities for additional runs of the stock assessment models should be located near enough to the meeting facilities so that STAT members do not miss out on discussions for the other stocks. These four species were reviewed at the same meeting because they are caught in a mixed species fishery and trends for one species may have an impact on the fishery for the others. Very little of the discussion focused on this issue. In addition, we often discussed modeling issues for one species that could have also been relevant to the other assessments. Having all of the STAT members there for these discussions would have been helpful.

Appendix 1. Bibliography

I. Current Draft Stock Assessments

A. Dover sole

1. The Status of Dover Sole off the U.S. West Coast in 2005, David B. Sampson, *DRAFT*
2. Control and data files for the preliminary base-run model, 2005 Dover sole assessment

B. Sablefish

1. Status of the Sablefish Resource off the Continental U.S. Pacific Coasts in 2005, Michael Schirripa and J.J. Colbert, *DRAFT*

2. Schirripa, M.J, and J.J. Colbert. 2005. Interannual changes in sablefish (*Anoplopoma fimbria*) recruitment in relation to oceanographic conditions within the California Current System. Fish. Oceanogr. 14:4, 1–12.

C. Longspine thornyhead

1. Stock Assessment and Status of Longspine Thornyhead (*Sebastolobus altivelis*) off California, Oregon and Washington in 2005, Gavin Fay, *DRAFT*

D. Shortspine thornyhead

1. Status and future prospects for the shortspine thornyhead resource in waters off Washington, Oregon and California as assessed in 2005, Owen Hamel, *DRAFT*

II. Background Materials

A. 2004 Workshop Reports

1. A Summary Report from The West Coast Groundfish Data Workshop held July 26-30, 2004 in Seattle, Washington. Northwest Fisheries Science Center. February 16, 2005.
2. A Summary Report from the Stock Assessment Modeling Workshop held October 25-29, 2004 at the Northwest Fisheries Science Center, Seattle, Washington. Northwest Fisheries Science Center, FRAM Division. March 16, 2005

B. Data Reports

1. Model-based estimates of abundance for 11 species from the NMFS Slope Surveys. Thomas E. Helser, Ian J. Stewart, Curt Whitmire, and Beth Horness. 2005.
2. Length and Age Composition Calculations for the NWFSC West Coast Survey of Groundfish Resources for the 2005 Assessment Season. Owen S. Hamel. April 29, 2005.
3. Standardized Catch Rates for the Deep-Water Complex. Jon Brodziak. 1997. (Note: Figures are currently not available in an electronic format but hard copies will be provided at the STAR Panel).

C. SS2 Documentation

1. Technical Description of the Stock Synthesis II Assessment Program. Version 1.17. Richard D. Methot. March 2005.
2. User Manual for the Assessment Program Stock Synthesis 2 (SS2), Model Version 1.17. Richard Methot. April 4, 2005.
3. PowerPoint Presentation: SYNTHESIS 2: Integrated Analysis of Fishery and Survey Size, Age, and Abundance Information for Stock Assessment. Richard Methot.
4. SS2 Model and Examples

D. STAR Panel Terms of Reference

1. Groundfish Stock Assessment and Review Process for 2005-2006. The Scientific and Statistical Committee (SSC) of the Pacific Fishery Management Council. 2005.

E. GAO Report

1. Pacific Groundfish: Continued Efforts Needed to Improve Reliability of Stock Assessments. United States General Accounting Office, Report to Congressional Requesters. June 2004.

III. Previous Stock Assessments and STAR Panel Reports

A. Dover Sole

1. Stock status of dover sole off the U.S. West Coast in 2000. D. Sampson and C. Wood. 2001.
2. Dover sole STAR Panel Report, 2001.

B. Sablefish

1. Status of the Sablefish resources off the continental U.S. Pacific Coast in 2002. Michael Schirripa. 2002. (Updated Assessment).
2. Review of the updated 2002 sablefish stock assessment (STAR Panel Report).
3. Status of the Sablefish Resource off the U.S. Pacific Coast in 2001. Michael Schirripa and Richard Methot, National Marine Fisheries Service. 2001.
4. Status of the Sablefish Resource off the U.S. Pacific Coast in 2001. Ray Hilborn, Juan L. Valero and Mark Maunder. School of Aquatic and Fishery Sciences, University of Washington. 2001.
5. Sablefish STAR Panel Report, 2001.

C. Longspine Thornyhead

1. Status of the thornyhead resources in 1997. J. B. Rogers, L.D. Jacobson, R. Lauth, J.N. Ianelli, and M.Wilkins. 1997. (Includes both shortspine and longspine thornyheads)

D. Shortspine Thornyhead

1. Stock status of shortspine thornyhead off the Pacific West Coast of the United States 2001. Kevin Piner and Richard Methot. 2001.

2. 2001 Shortspine thornyhead STAR panel Report

Additional materials:

Dick, E.J. 2004. Beyond ‘lognormal versus gamma’: discrimination among error distributions for generalized linear models. Fisheries Research. 70: 351–366.

Helser, T.E, Punt, A.E., and Methot, R.D., 2004. A generalized linear mixed model analysis of a multi-vessel fishery resource survey. Fisheries Research. 70: 251–264.

Lauth, R. R., J. Ianelli and W. W. Wakefield 2004. Estimating the size selectivity and catching efficiency of a survey bottom trawl for thornyheads, *Sebastolobus spp*. Using a towed video camera sled. Fisheries Research 70: 27–37.

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

Stefánsson, G., 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES J. Mar. Sci. 53: 577–588.

Appendix 2: Copy of the statement of work

Statement of Work
Consulting Agreement between the University of Miami and Dr. Stephen Smith

May 24th, 2005

General

External, independent review of West Coast groundfish stock assessments is an essential part of the STAR panel process. The stock assessments will provide the basis for the management of the Dover sole, sablefish, shortspine thornyhead and longspine thornyhead stock assessments.

The consultants will participate in the Stock Assessment and Review (STAR) Panel of the Pacific Fishery Management Council (PFMC) for the review of the Dover sole, sablefish, shortspine thornyhead and longspine thornyhead stock assessments. The consultant should have expertise in fish population dynamics with experience in the integrated analysis type of modeling approach, using age-and size-structured models, use of MCMC to develop confidence intervals, and use of Generalized Linear Models to process survey and logbook data for use in assessment models.

Documents to be provided to the consultants prior to the STAR Panel meeting include:

- Current drafts of the Dover sole, sablefish, shortspine and longspine thornyhead stock assessments;
- Most recent previous stock assessments for Dover sole, sablefish, shortspine and longspine thornyhead;
- An electronic copy of the data, the parameters, and the model used for the assessments (if requested by reviewer);
- The Terms of Reference for the Stock Assessment and STAR Panel Process for 2005-2006;
- Summary reports from the West Coast Groundfish data and modeling workshops held in 2004;
- Stock Synthesis 2 (SS2) Documentation; and
- Additional supporting documents as available.

Specifics

Consultant's duties should not exceed a maximum total of 14 days: several days prior to the meeting for document review; the 5-day meeting; and several days following the meeting to complete the written report. The report is to be based on the consultant's findings, and no consensus report shall be accepted.

The consultant's tasks consist of the following:

1) Become familiar with the draft stock assessments and background materials.
2) Actively participate in the STAR Panel to be held in Newport, Oregon from June 20-24, 2005. . *Participants are strongly encouraged to voice all comments during the STAR Panel so the assessment teams can address the comments during the Panel meeting.*
3) Comment on the primary sources of uncertainty in the assessment.
4) Comment on the strengths and weaknesses of current approaches.
5) Recommend alternative model configurations or formulations as appropriate during the STAR panel.
6) Complete a final report after the completion of the STAR Panel meeting.
7) No later than July 8, 2005, submit a written report consisting of the findings, analysis, and conclusions (see Annex I for further details), addressed to the "University of Miami Independent System for Peer Review," and sent to Dr. David Die, via e-mail to ddie@rsmas.miami.edu, and to Mr. Manoj Shivlani, via e-mail to mshivlani@rsmas.miami.edu.