

DRAFT SUMMARY MINUTES
Groundfish Stock Assessment Review Workshop

Pacific Fishery Management Council
Sheraton Portland Airport Hotel
Columbian A Room
8235 NE Airport Way
Portland, OR 97220
503-281-2500

January 13, 2006

FRIDAY, JANUARY 13, 2006 – 8 A.M.

Attendees:

Dr. Elizabeth Clarke, National Marine Fisheries Service Northwest Fisheries Science Center
Dr. Martin Dorn, National Marine Fisheries Service Alaska Fisheries Science Center, Scientific and Statistical Committee
Dr. Steve Ralston, National Marine Fisheries Service Southwest Fisheries Science Center, Scientific and Statistical Committee
Dr. Alec MacCall, National Marine Fisheries Service Southwest Fisheries Science Center
Dr. Jim Hastie, National Marine Fisheries Service Northwest Fisheries Science Center
Dr. David Sampson, Oregon State University, Scientific and Statistical Committee
Dr. Richard Methot, National Marine Fisheries Service
Mr. Jason Cope, University of Washington
Dr. Mark Maunder, International Tropical Tuna Commission, Quantitative Resource Assessment LLC
Mr. Ian Stewart, National Marine Fisheries Service Northwest Fisheries Science Center
Dr. Owen Hamel, National Marine Fisheries Service Northwest Fisheries Science Center, Scientific and Statistical Committee
Dr. Robert Mohn, Department of Fisheries and Oceans, Center of Independent Experts
Mr. Tom Jagielo, Washington Department of Fish and Wildlife, Scientific and Statistical Committee
Mr. Guy Fleischer, National Marine Fisheries Service Northwest Fisheries Science Center
Ms. Meisha Key, California Department of Fish and Game
Ms. Michele Culver, Washington Department of Fish and Wildlife, Groundfish Management Team
Mr. Brian Culver, Washington Department of Fish and Wildlife, Groundfish Management Team
Mr. Mark Saelens, Oregon Department of Fish and Wildlife, Groundfish Management Team
Dr. Steve Berkeley, University of California Santa Cruz, Scientific and Statistical Committee
Dr. Michael Schirripa, National Marine Fisheries Service Northwest Fisheries Science Center
Mr. Tom Ghio, Groundfish Advisory SubPanel
Dr. John Field, National Marine Fisheries Service Southwest Fisheries Science Center, Groundfish Management Team
Mr. Curt Melcher, Oregon Department of Fish and Wildlife
Mr. Rishi Sharma, Columbia River Intertribal Fish Commission
Mr. Henry Yuen, United States Fish and Wildlife Service

Mr. Steve Joner, Makah Tribe
Mr. Hap Leon, Makah Tribe
Mr. Rob Jones, Northwest Indian Fish Commission, Groundfish Management Team
Mr. Kelly Barnett, Independent Fish Filleter, Bay City, Oregon
Ms. Stacey Miller, National Marine Fisheries Service Northwest Fisheries Science Center
Mr. Pete Leipzig, Fishermen's Marketing Association
Mr. Brad Pettinger, Oregon Trawl Commission
Mr. Dan Waldeck, Pacific Whiting Conservation Cooperative
Mr. Steve Theberge, Oregon Sea Grant
Dr. Donald McIsaac, Pacific Fishery Management Council
Mr. Mike Burner, Pacific Fishery Management Council
Mr. John DeVore, Pacific Fishery Management Council

A. Administrative Matters

1. Roll Call, Introductions, Announcements, etc.

Dr. McIsaac called the meeting to order at 8:20 a.m. A round of introductions was done.

2. Opening Remarks and Agenda Overview

Mr. DeVore reviewed the agenda. He explained these minutes would help inform the Council process to plan the next suite of stock assessments and the review of those new stock assessments. The Council will consider recommendations from the workshop participants at the March Council meeting. Final adoption of the next round of assessments; new Terms of Reference for stock assessments and assessment reviews; new Terms of Reference for Rebuilding Analyses; and all the other elements of the Council assessment process will occur at the April Council meeting.

B. Perspectives on 2005 Stock Assessment Process

1. Groundfish Management Team (GMT) Perspective

Mr. Saelens provided the GMT perspective. The general thought was 23 assessments were too many done in too brief a period. About 10-15 assessments per cycle seemed more reasonable. The GMT also believed there were too many assessments reviewed per Stock Assessment Review (STAR) panel. In many instances, a base model was not decided until late in the week. It might help to have the most contentious and/or complicated assessments scheduled early in the process. Every effort was made to get the information out to the public as early as possible. The GMT recommends a debriefing by the Stock Assessment Team (STAT) after a STAR panel recommends an assessment. The GMT wants more specific advice from STAR panels and the Scientific and Statistical Committee (SSC) on how the science should be applied to management decision-making. There was also concern that not all assessments had the requisite management estimates and other details mandated by the Terms of Reference. This includes lack of timely delivery of STAR panel reports when an assessment is first considered. Most of the executive summaries in assessments were clear and concise, but improvements can be made, particularly with respect to what models and estimates should be used to formulate management advice. The GMT recommends a management trigger to decide which assessments should be done next. The importance to management or risks to stocks of overexploitation are example triggers for making

this decision. The GMT agrees early decision-making on the next cycle or more of assessments should be done for contributors to prepare data used in assessments. The GMT recommends the next process be set up to get adequate interactions between STAT teams and Council advisors. Dr. Dorn asked if it would help if STAR Panel chairs attended a GMT meeting. Mr. Saelens said that would help, but it was hard with filled agendas to carve out that time. Perhaps some triage to get such interactions for more complicated or contentious assessments is the answer. Need to do a better job planning sampling priorities for use in assessments. This was a time-honored process that has slipped in recent years. The GMT recommends a greater amount of biological sampling needs to occur in the at-sea observer program. The GMT wants a greater role in the STAR panel process. The GMT could help set a probable range of optimum yields (OYs) earlier in the assessment process (i.e., during the STAR panel). The GMT is concerned with the range of methods used to account for total mortality in assessments. A more consistent approach is requested. Ms. Culver expressed the opinion that the number of assessments done last year (23) was not necessarily too many, if the process was changed to accommodate that many (i.e., fewer assessments reviewed per STAR panel). The GMT wants to set up future processes to allow new stocks to be assessed. All overfished species are assessed each cycle and other important stocks are frequently assessed as well, which limits the number of new assessments. The plan needs to incorporate the quality of assessment data as well as risks to the stock when deciding which stocks get assessed. Mr. Culver said better planning on the data going into assessments would benefit the process. He cited the problems with the petrale sole assessment where the STAT Team was not aware of critical data gaps until the STAR panel. Contributing agencies need to tune in to the pre-assessment data workshop. Dr. Clarke commented that the observer program is collecting some ageing data. They are trying to balance the amount of time dedicated to at-sea catch and biological sampling. Dr. Mohn said he was surprised at the lack of risk plots in the Council's assessment process. Mr. Saelens said he thought that should be the fundamental structure of an assessment decision-table. (Currently, decision tables are used to address assessment modeling uncertainty) Perhaps both treatments are needed: one decision table to depict assessment uncertainty and one table depicting risk of alternative mortality schedules using the most plausible base model in the assessment.

2. Groundfish Advisory SubPanel (GAP) Perspective

Mr. Ghio provided the GAP perspective. Originally, STAR panels were set up to solicit industry input and instill industry confidence in the assessment process. Now, the GAP representative to a STAR panel has a much diluted role. There is also a critical need to have a STAT team member or the STAR panel chair interact with the GAP to answer questions on assessments. The GAP needs a better understanding of the assessment details to make informed recommendations on OYs. In general, he agreed with the comments and recommendations of the GMT as expressed by Mr. Saelens. Mr. Ghio thought one of the priorities in the assessment decision-making process is to assess the more constraining and/or valuable species more frequently. STAR panels spent the majority of time on modeling approaches rather than the quality of input data or the assessment result.

3. SSC Perspective

Dr. Dorn provided the SSC perspective and requested other SSC members in attendance to chime in. The SSC were the architects in this process and did warn the quantity of assessments would compromise the quality. This did occur to some degree; however, the overall quality of assessments given the magnitude of the task is laudable. Now we should think about how to make the process more efficient by doing fewer things better. Shifting more stock assessments into an update mode is one way to achieve this goal. Especially now that the Stock Synthesis 2 (SS2) model is tried and true, there should be more stability in the next cycle. Thoughtful planning on which new assessments should be done will help keep future problems to a minimum. Dr. Ralston agreed one of the major changes to the process last year was to have SSC members chair STAR panels. This provided continuity despite the extra workload. This was a good idea and should be continued. Dr. Dorn thought one problem was the STAR chair became too wedded to the STAR panel recommendations. There should be more sensitivity to outside views of SSC members and other advisors by the STAR chairs after a STAR panel has finished their business. Mr. Jagielo emphasized the problem with the process last year was everyone was working in a new modeling environment (SS2). However, he gave kudos to Dr. Methot for helping everyone understand the complexities of SS2. Dr. Berkeley recommended some greater thought in what kinds of data should be incorporated in an assessment. More consideration of the age structure of the spawning stock and genetic structures should be incorporated more thoroughly in assessments. Ecosystem-based principles needed to be considered as well. More complex spatial management issues need to be considered more thoroughly. Perhaps trophic relationships should be more thoroughly considered as well. One possible improvement might be to do multi-species assessments instead of single species assessments. The scientific community is currently grappling with these issues. Dr. Mohn thought some of the Council's data issues compromise current single-stock assessments. There should be more thought on the biological data sampling and some of the shortcomings of current data (i.e., lack of ageing structures, etc.). Dr. MacCall thought these issues should be addressed during the assessment "off-year". Dr. Field said North Pacific scientists are starting to incorporate food web/trophic relationships in assessment and cited recent pollock assessments. Dr. Berkeley said spatial considerations are critical to avoid localized depletion and other problems. Dr. Dorn said there is a current mandate to evaluate ecological, spatial issues. There is a problem with the lack of informative data that are useful. Dr. Clarke said the Council also needs to plan on how they use ecosystem-based information in their management decisions. Dr. Ralston said there is currently tension between doing updates and incorporating ecosystem considerations in assessment. Such new assessments require more thought and review and will compromise the number of assessments that can be done in a cycle. Dr. Hamel said it will take time to decide how such new information will be incorporated in assessment and used in management decision-making.

4. Northwest Fisheries Science Center (NWFSC) Perspective

Dr. Clarke provided the NMFS perspective and had a Powerpoint presentation to emphasize her points. She also handed out a summary of reviewers' comments to the 2005 STAR process. She reviewed the previous process from the data workshop through the variety of STAR panels. Each STAR panel was chaired by an SSC member and the number of panel members was N (number of assessments) plus 1. There was a single CIE reviewer (Dr. Mohn) who attended each STAR panel except for the hake STAR panel. SS2 was used in most cases. The NWFSC requested summary items for the executive summary of each assessment (new mandate). The NWFSC requested assessment authors list all their intended input data after the Data Workshop. This was done to prepare data contributors, but was only partially successful. Another new item

was the “sweep-up” or “mop-up” STAR panel, which was useful (3 assessments were referred to this panel).

Some shortcomings to the process: Several assessments were incomplete coming into the STAR panel or were distributed later than two weeks prior to the STAR panel. There was not enough participation in the Data Workshop. This was a new workshop and more people are expected to tune in to the next workshop. There was not enough utilization of assessment authors’ list of data needs. There were too many assessments done last year and too many assessments reviewed at each panel. There were no species-specific data workshops prior to STAR panels. There were no species-specific presentations to the Council family after the STAR panel.

We need to collectively decide the goal of the assessment process. We should be explicit that the goal is not to assess every groundfish species managed under the FMP. This is neither attainable nor critical to the process. She underscored the value of Stacey Miller’s coordination efforts. She solicited comments from the groups on how to make the logistics more efficient. Dr. Ralston requested more dialogue between the NWFSC and the Southwest Fisheries Science Center (SWFSC). Dr. Clarke said the SSC was the group the NWFSC coordinated with. Dr. Ralston said the SWFSC is doing more planning regarding sampling and assessments. Dr. Clarke agreed more coordination with the SWFSC was a good idea. Dr. Methot said one of the overarching objectives implied in the law and policy is to assess all species. Perhaps a multi-species approach should be more carefully considered. What do we need to say to defer a species’ assessment? What is the probability we are avoiding overfishing in the current and future management regime if an assessment is not done? These are important questions that need to be answered in the process. Dr. MacCall said there seemed to be many entities involved that appeared to be in charge (NWFSC, SSC, Council, and Council staff). Dr. Clarke said the Council is in charge of this process according to FACA rules. The Council delegates science advice and direction to the SSC and logistic support to the NWFSC. Her perspective is that many entities chipped in to help move the process along. Dr. MacCall requested more formal direction from the Council. Dr. Methot said, while ownership of the process is with the Council, there is the issue of corporate ownership of the assessment itself. NMFS relies on the Council review process to decide whether an assessment is the best available science despite the fact that the agency does an independent review of the quality of the assessment and the review. Dr. Clarke said the NWFSC works with the Council to ensure the Council science review process satisfies OMB rules and policies. For any assessment coming from the NWFSC, she signs a letter validating that it complies with the OMB circular. Dr. Dorn said the modeling environment of SS2 is now locked into the Council process. Perhaps different types of assessments and assessment models should be considered. A lot of data is left unexamined. Dr. Methot said there is nothing in the Terms of Reference mandating SS2. Dr. Dorn stated implicit in the current process is the need to fit the data to an assessment model. Returning to Dr. Methot’s recommendation to address all FMP species is to use what you do know in a NEPA document or a multi-species assessment to decide management risk. For instance, identify species with de minimus exploitation. Dr. Schirripa thought that to do this, more flexible rules/guidelines are needed for determining if a stock is overfished or experiencing overfishing. Dr. Methot said there are a variety of processes used nationwide to decide management risk. It is better to gain your flexibility in methods for determining stock status and risk rather than changing the rules.

5. Council Perspective

Dr. McIsaac polled Council members on their perspectives. He talked with Frank Warrens, Don Hansen, Dave Ortmann, and Bob Alverson; and also received emails from Patty Burke, Rod Moore, and Marija Vojkovich. General comments: 23 assessments were too many (with current resources), there were too many full assessments, the roles of GAP and GMT representatives on STAR panels needs to be formalized, 4 assessments per STAR panel were too many and 3 may be pushing it, there was not enough debriefing on individual assessments (many Council members were not comfortable with the advice, or lack of advice, given), more specific management advice is needed (e.g., how should the kelp greenling assessment be used for management?), consistent and useful decision tables are needed, timely delivery of stock assessments and STAR reports was lacking in some instances, more resources/funding for this process is needed, guidelines are needed to prioritize assessments, and many Council member gave their compliments for achieving such an ambitious goal of doing 23 assessments.

Mr. Saelens complimented staff for their roles in the process. He encouraged assessment authors to date-stamp all their documents to maintain version control.

C. Improving the Stock Assessment Process

1. Pre-Assessment Planning
 - a. What Worked and What Didn't in 2004-2005
 - b. Recommended Improvements for 2006-2007

Dr. Clarke said one consistent recommendation from reviewers was better reviews depended on scheduling fewer assessment reviews per STAR panel. Addressing this requires either fewer assessments be done, more reviewers dedicated to the process, or fewer reviewers per panel. Most reviewers thought 2 assessments per panel was ideal for a 1-week panel with the potential of adding a third assessment if they are not too complicated. Some reviewers and assessment authors stated their workload was too high. Dr. Sampson said incorporating the new SS2 model during 2005 added to the workload because of the steep learning curve. Dr. Clarke said another factor is the delivery of assessment input data. For instance, trawl survey data from the previous summer and fall are delivered in February. Some NWFSC recommendations: try to schedule only 2 assessments per panel, plan for a maximum of 8-10 full assessments, not all assessments need to be done each cycle, clear criteria should be developed for determining priorities for full assessments, updates should be reviewed by the SSC only. Dr. Dorn thought a separate meeting of the SSC Groundfish Subcommittee to review all updated assessments might be advisable. The review itself should be expedited and could be done early in the process. Another idea is to review 2 full assessments plus 1 updated assessment per STAR panel. Dr. Mohn thought that process improvement would work; however, there was a problem with some "updated" assessments coming in that were not really updates (i.e., yelloweye). Dr. Mohn recommended reviewing "benchmark" data-poor assessments early in the process. Much can be learned reviewing such assessments that can be applied in later assessment reviews. Mr. Ghio said a disciplined approach in the assessment review is needed to ensure assessments comply with the Terms of Reference. Mr. Burner said scheduling an SSC review of updated assessments was and could continue to be problematic given SSC members' heavy workload. Mr. Saelens thought updated assessments could be done and reviewed in the off-year. Dr. McIsaac said if the intent is to use these assessments in the next management cycle, then it is hard to defend that decision-making uses the best available science if a reviewed assessment is sitting on the shelf. The potential benefits of multi-year management are eroded in this case and the management regime

is vulnerable to legal challenge. Ms. Culver asked if there was a requirement to do a full assessment of overfished stocks every cycle. The mandate is to review rebuilding plans at least once every other year and a full assessment, or even an updated assessment, is not necessary. A data review may be sufficient to decide whether catches are staying within OYs; however, it may be advisable to at least do an updated assessment every cycle for a stock under rebuilding. Dr. McIsaac encouraged brainstorming on a policy for this and let the Council decide a policy later after receiving recommendations. Dr. MacCall recommended putting into the Terms of Reference a requirement for STAR panels to specify whether an assessment is sufficiently developed to do an update in the next cycle.

Dr. Clarke discussed recommended criteria for determining assessment priorities and whether they should be full assessments for the next cycle. Criteria include assessments not currently done using SS2, new stock assessments, etc. for deciding whether an assessment should be a full one. Dr. Hastie said NMFS Headquarters has also determined if an assessment hasn't been done within five years, it is considered out of date.

Dr. Ralston recommended pre-assessment planning within NMFS between the NWFSC and the SWFSC before a recommendation is brought forward into the Council process. Data needs and resource capacity needs to be internally deliberated to decide what can be done. Dr. Clarke thought the first step is to develop criteria for deciding the list of stocks that should be assessed.

Dr. Clarke recommended a pre-STAR data workshop to prepare for assessments. Dr. Dorn thought it would be more useful if there were multiple single-species workshops to bring the right people into the process and deliberate assessment data needs in more detail. Mr. Culver recommended a sampling meeting earlier in the process to prioritize preparation of assessment data. Dr. Methot thought workshops that are subject oriented, such as a trawl survey workshop, would be most efficient. Such workshops would be useful for a host of stock assessments.

A discussion of how to balance the assessment workload with a mix of full and updated assessments ensued. There is a dynamic tension between allowing creativity and innovation in an assessment and engineering a stable process by planning for updated assessments that do not need such a comprehensive review as a full assessment. With the limited resources available, this will remain a conflict.

Dr. Sampson asked if we need a modeling workshop for the next cycle. There were other ideas for workshops beyond those already recommended for this year. Issues, such as how to address model uncertainty, require a lot of careful planning and preparation. Workshops are expensive and need to be carefully planned and prioritized. Some workshops are generic (i.e., the data workshop) and others are specific to one issue (i.e., the contemplated juvenile survey workshop).

2. Stock Assessment Reviews

- a. What Worked and What Didn't in 2004-2005
- b. Recommended Improvements for 2006-2007

Dr. Mohn provided his perspectives on the 2005 STAR process. He thought the process would benefit from more pre-review vetting of data issues. Assessments need more documentation of input data and better descriptions of methodologies. In some cases, which were rare but serious, pre-review assessment drafts were incomplete. In his opinion, there was inadequate contemplation of data (and this was more common). More interaction with fishermen would benefit the sensibility and quality of input data. There was too little context of why assessments

were done the way they were. The previous STAR reports with lists of recommended improvements were helpful, but needs to be emphasized. There was too little synthesis of data in pre-review assessments. Requests for re-runs/alternate model runs varied from panel to panel and were not always clearly articulated to STAT Teams – this could be improved. The STAR review schedule was too ambitious- review fewer assessments to improve quality of assessment reviews. Updates should stay as updates- inject more discipline in the review process. The roles of STAT Teams vs. STAR members created a natural tension that was hard for STAR chairs to fix. The SSC, having such a strong role in the STAR process, was mostly positive despite their lack of independence in the process. The choice of a chair for the STAR panel is perhaps more critical than reviewers because there are a number (3-4) of reviewers to help one another, but the chair has no back-up. Communication in the process could be improved. Not all STAT teams attended all the Council, SSC and other meetings that were integral in the process. Treatment of uncertainty varied greatly from panel to panel. Determining model plausibility was done in a somewhat ad hoc fashion. More thought could be given to this. One recommendation might be to do a real-time meta summary of important parameters, such as M (natural mortality rate), h (stock-recruitment steepness), and q (catchability), as well as the recruitment time series by species. This could help identify outliers and potentially aid in model selection. Synthesizing estimated vs. assumed parameters and listing/plotting these parameters would be informative. Another recommendation is to plot the difference in model results pre- and post-STAR. There are competing philosophies on whether the STAT Team prepares the draft assessment to a minimal level anticipating that the STAR panel will change it or defends a more complete draft assessment at a STAR panel. Dr. Clarke said that during 2005 UW and NMFS staff did a lot of vetting of assessments prior to STAR panel review and that helped improve the quality of pre-STAR draft assessments.

Logistic improvements: provide a LAN router/printer at each STAR panel, more attendance of STAT members at more STAR panels, off-season benchmark assessments and reviews could help set the stage for the next round of assessments and reviews (benchmark assessments can be characterized as a prototype and can be done by committee), and a second STAR reviewer who goes to all STAR panels to provide more continuity. There should be more attention paid to the use of priors in assessments. A workshop on the use of CPUE data may be particularly useful in this next assessment process. Dr. Clarke remarked that more and better input from the GAP and GMT on this subject would be useful. Dr. Mohn said the non-linearity of CPUE trends from various surveys is a problem to resolve. Dr. Maunder thought it could be helpful to appoint a data expert on a particular dataset used as an index (i.e., Dr. MacCall's expertise in using recreational CPUE). Dr. MacCall said the process evolved from more ad hoc assessments, which led to the process of dismissing data wholesale, as is the current practice. He disagreed with the process of limiting input data. Others disagreed and championed thoughtful analysis of data before deciding whether they were useful for an assessment. Dr. Mohn said the important distinction is whether the data are used to inform the assessment versus their use in tuning the assessment. Data used to tune an assessment need to be carefully discriminated. Dr. Dorn said part of this discrimination is the quality of the analysis used to create the index. Dr. McIsaac asked about the comment that some assessments were incomplete coming into the STAR panels last year. He wanted to know how frequently that occurred and Dr. Mohn said it was rare. Dr. Clarke said about 25% of the assessments were delivered late (not within two weeks prior to the STAR panel).

Dr. Sampson gave the SSC's perspective on the review process, although he said he was not able to poll SSC members and these comments should be considered his own. He thought lessons learned at the first STAR panel for flatfish were useful in subsequent panels. He recommended

capturing and disseminating these lessons. There needs to be better follow-up on recommendations for improving assessments from previous STAR panels. One problem that arose last year is that data issues would crop up during a STAR panel that led to snap decisions. There was inconsistent treatment of common issues, such as treatment of residuals, effective sample sizes, outliers, spatial structure, etc. He recommended that assessment scientists should not serve on more than one of the STATs being reviewed during a STAR panel. Dr. Clarke said, while they know the lead authors for an assessment ahead of time, they were not informed of all STAT team members prior to setting up the STAR panel. More interaction with industry members early in the process prior to a STAR panel would be helpful. Dr. Sampson thought the process loses some of its independence by having SSC members serving on STAT teams, STAR panels, and on the SSC during the final review phase. Dr. Ralston said part of the issue is deciding if a STAR panel needs to be an independent review body or a peer review body. Dr. Sampson said, while it may be appropriate to have SSC members chair STAR panels, there is a cost to the independence of the SSC members in their respective roles. Dr. Ralston pointed out that, in one circumstance, the SSC rejected an assessment (and sent it to the mop-up STAR panel) after the first STAR panel recommended it. This indicates the SSC maintained some level of independence. Dr. McIsaac asked how often SSC members voted within the SSC on their own assessment and was told those members recused themselves from the vote. Dr. Dorn thought this should be formalized in the Terms of Reference and that these members should not write the SSC report on their own assessments. Dr. MacCall said we should not make SSC members totally independent of the review process since the process would lose that expertise. Mr. Ghio asked if the SSC review step should be a two-meeting process and was told there was not enough time in the process to do this (at least there was not enough time in last year's process). Ms. Culver thought the final review step could be a joint SSC-GMT deliberation.

Dr. Clarke provided the NMFS perspective and recommended improvements regarding the review process. She recommended distributing STAR panels coastwide as was done last year. There should be N plus one reviewers at a STAR panel with a maximum of three reviewers. There should be at least one reviewer independent of NMFS and the process at each STAR panel (CIE reviewers are not available for all panels). Dr. Ralston pointed out that some NMFS personnel may be independent of the process and should be considered in the STAR process as an independent reviewer. Dr. Sampson thought it was problematic to limit the panel to three reviewers as a maximum. SSC members should continue to serve as STAR chairs, but should rotate through the stock reviews. There should be strict adherence to Terms of Reference and update criteria. Should there be a rapporteur assigned for all STAR panels in a cycle? Rapporteurs should produce real-time assessment summary tables, which should be produced at each panel to track changes in the assessment as the review progresses. The rapporteur would also track requests by the STAR panel and the responses made by the STAT. The rapporteur duties should not fall to the chair or the CIE reviewer. These notes and a summary of changes to the assessment could be posted to a web site so other STAT teams and STAR panel members can better understand how each panel addressed review issues. Dr. Mohn did not think it was too onerous to be a rapporteur at a STAR panel and at the same time effectively critique the assessment. Dr. Hastie thought STAR panel members can share rapporteur duties as is current practice.

Assessments need to be complete coming into a STAR panel. These draft assessments need to be internally reviewed prior to delivery to a STAR panel (if not just to make them more readable). Decision tables should be incorporated in pre-STAR drafts and should be standardized. Executive summaries should be complete and included in the pre-STAR draft. There needs to be a standard minimum set of diagnostics produced for each model run. Each

assessment needs an explicit section responding to recommendations for improvement from previous STAR panels. Dr. Sampson pointed out this used to be in the Terms of Reference but was dropped recently. A review of available information from Canadian and Alaska assessments should be included in each assessment. Finally, maps showing the geographic scope of the assessment need to be included.

Mr. Ghio provided some of the GAP and industry perspective in the assessment review process. There is general distrust by industry of the process. To rectify this, industry needs to have direct input in the process prior to the STAR panel. Dr. Clarke said the intent of listing data to be used in each assessment early in the process was to solicit industry feedback, but the approach was not effective. Ms. Key said she reached out to industry to prepare for the gopher rockfish assessment and thought that helped the quality of her assessment. Individual authors approached this type of feedback differently. Dr. Clarke thought this might be difficult to formalize in the Terms of Reference. Dr. MacCall recommended that GMT and GAP members should be formally recognized and incorporated in the STAR process. Dr. Sampson thought this would work under an N + 3 process, but not an N + 1 process. Dr. Sampson thought coordination with the port liaison project (or other source of funding for fishermen) would benefit assessments by providing ideas for data inputs that might not be obvious to an assessment author. Dr. Hastie said older versions of the pre-assessment workshop were more hands-on with lots of industry input. This would be a good way to reintegrate industry early in the assessment planning process.

3. Scientific and Statistical Committee Reviews
 - a. What Worked and What Didn't in 2004-2005
 - b. Recommended Improvements for 2006-2007

Dr. Dorn provided a review of last year's SSC review process. All assessments were reviewed in two SSC meetings with each STAR chair leading the discussion on the assessments reviewed at their panel. The chief concerns were having too many assessments in the cycle and too many assessments reviewed at each panel, which compromised the quality of the review. This also compromised the SSC review step. For example, new assessment elements, such as the environmental index in the sablefish assessment and the canary assessment, did not receive adequate attention by the SSC at their meetings. The SSC needs to be more sensitive to STAR chairs defending the STAR reviews they chaired. Dr. Sampson thought it might be better if an SSC member who did not chair a particular STAR panel lead the discussion in the SSC meeting. Dr. Clarke wanted a better definition of the SSC's role in the review process. Drs. Sampson and Dorn thought the SSC's role should be review of the assessment, not reviewing the STAR panel report. Dr. McIsaac said the SSC's role is to recommend the best available science to the Council and needs to critically review all aspects of an assessment. Mr. Burner said the SSC's approach at their review last year was limited by the time they could allot to the review. He was not sure an SSC member could find the time to review and report on an assessment if that SSC member had not been present at the assessment's STAR review. Drs. Sampson and Ralston thought the extra cost of an SSC's member time is real, but independence of the SSC could be maintained if the SSC member reporting to the SSC on an assessment was not the person who had chaired the STAR for that assessment. Dr. Methot wondered if there was any opportunity to address SSC concerns during an SSC meeting and Dr. Hamel said the mechanism for further modeling to resolve issues is to send the assessment to a mop-up panel. Otherwise, the SSC decision is either approve or reject the assessment. Mr. Ghio reiterated the need to have the GAP and GMT review a post-STAR assessment as well. Dr. Ralston thought a good mechanism would be to have the independent SSC reviewer solicit GAP and GMT feedback during their

review. Dr. Dorn said, in the past, there were formal presentations of assessment results to the Council family. Such a debriefing was beneficial. Dr. Field thought subcommittees from the GMT and GAP could attend these debriefings to manage workload. Mr. Ghio said the main issue with the GAP is erosion of confidence in the assessment process due to lack of interaction with STAT teams and a loss of institutional knowledge within the GAP itself. Staggering the final Council review step across more meetings would help, as would reducing the number of assessments done. Mr. DeVore explained one GMT recommendation was to improve SSC advice on how assessment results should be used. Some of the SSC statements were somewhat vague in that regard. Dr. Hastie said another aspect of GMT deliberations on recommending OYs based on a new assessment are that there is reticence to recommend alternative model results or recommending an OY alternative for analysis that is above the acceptable biological catch (ABC) from the base model in the assessment. Dr. Field said there was discomfort in the GMT recommending specifications from alternative models since there was thought these models did not represent the best available science. Dr. Ralston offered an approach where one decision table is prepared to address model uncertainty and another decision table is prepared to address statistical uncertainty within the base model. Dr. Methot thought the approach taken in the most recent canary assessment, blending the uncertainty of equally plausible models, to be fertile ground for an approach. Alternative models that are not considered as plausible could also be blended with a weighted approach based on probability distributions. If this approach was conceptually accepted, the modeling details could be worked out.

Dr. MacCall thought some consideration to trade assessment authorship with outside entities might help. That is, west coast assessment scientists could do some east coast assessments and vice versa.

D. Terms of Reference

1. Review the “Terms of Reference for Groundfish Stock Assessment and Review Process for 2005-2006” and Provide Recommended Edits
2. Review the “Terms of Reference for Groundfish Rebuilding Analyses” and Provide Recommended Edits

The workshop participants discussed the process for modifying the Terms of Reference. This will be done at the March and April Council meetings and will be informed by advice from participants at this workshop and other advisors to the Council. The Council will adopt a final Terms of Reference in April.

Dr. Hastie asked if the Terms of Reference would be cycle-specific or more generic to the process. Dr. Ralston said these documents have evolved over time and would likely continue to change. Dr. Clarke asked how comments need to be provided- redline/strikeout or general comments? Dr. Ralston said the former vehicle is more specific and clear. Dr. Mohn thought the existing Terms of Reference for stock assessments was quite useful and the rebuilding Terms of Reference less so.

Dr. Ralston reviewed the elements of the current stock assessment Terms of Reference. The roles and responsibilities of various entities are explained followed by stock assessment priorities, terms for STAR panels and their meetings, suggested template for STAR reports, and terms for STAT teams. The appendices go into further detail on what needs to be included in assessments and STAR reports.

Dr. McIsaac asked what the penalty should be if Terms of Reference are violated? Some items should be truly mandated, such as timely delivery of STAR reports (according to some participants), where failure to provide these elements causes the assessment to be rejected. Other omissions are less egregious and should not result in such a harsh penalty. Dr. Ralston thought each violation should be judged specifically to each case- it's hard to draw lines in the sand. Dr. Methot said there were numerous instances of late delivery of critical assessment data. This compromised the timely delivery of assessments and the following STAR and Council process. He recommended a firewall on considering such late data. Ms. Key said another issue was the evolution of the SS2 model during the process. Many stated that many of the SS2 changes were made at the request of assessment scientists to Dr. Methot to fix certain aspects of the model. Dr. Clarke said the mop-up panel should be used to fix assessment problems identified during the review process, not to incorporate data arriving late in the process. Dr. McIsaac said it will be hard to list all possible offenses and requisite penalties in a Terms of Reference, but a process where rules are established regarding how and who decides if an omission is critical enough to reject an assessment could be considered for the Terms of Reference. Dr. Clarke said many of the fixes are on the front end with better data and modeling workshops and more time for assessment authors to do their work. While this won't solve all the problems encountered with late data delivery, it will help. Dr. MacCall said a schedule of deliverables would be helpful.

Dr. Ralston then reviewed the rebuilding analyses Terms of Reference. He agreed with Dr. Mohn's assessment that this Terms of Reference was not as useful. This Terms of Reference needs to incorporate an evaluation of existing rebuilding plans and revision rules if rebuilding progress is lagging or ahead of schedule. An ad hoc process was used in 2005 to have additional rebuilding runs done to evaluate rebuilding progress. Given uncertain outcomes in current litigation, Magnuson-Stevens Act (MSA) reauthorization, and potential changes to National Standard 1 (NS1) guidelines, it may not be useful to modify this Terms of Reference by April.

Dr. Clarke encouraged folks to send a list of candidate species for the next cycle to her. She would compile this list for the March briefing book. Dr. Hastie encouraged folks to look beyond the upcoming cycle when recommending which assessments are done next. Dr. Ralston encouraged folks to identify which assessments should be full assessments and which assessments should be updates. Dr. Hastie recommended folks provide the rationale for their recommendations and judge whether the draft NWFSC list has any fatal flaws. Dr. Clarke asked for a prioritized list (i.e., the top eight stocks for assessment). Dr. Dorn asked whether multi-species data summaries should be done.

Summary of Workshop Participants' Recommendations

Improving the Stock Assessment Process

Pre-Assessment Planning

- Fewer assessments (than the 23 done in 2005) should be done per cycle- consider a maximum of 10-15 full and updated assessments. Limit the number of full assessments to a maximum of 8-10.
- More pre-review vetting of data issues. Consider scheduling either subject-oriented (i.e., trawl surveys, CPUE indices) or species-specific workshops in the “off-year” prior to doing assessments.
- More interaction with fishermen in planning an assessment would benefit the sensibility and quality of input data.
- Develop guidelines for prioritizing full stock assessments in a cycle. Such guidelines include: the stock’s importance to management, relative risk of overexploitation, whether the stock has recently been assessed (NMFS Headquarters considers an assessment older than five years to be out of date), whether the most recent assessment uses the most up-to-date model (i.e., SS2), the quality of available data for that species, etc.
- Develop a schedule of deliverables with deadlines when planning an assessment.
- Identify which assessments should be full and which should be updates when recommending an assessment for the next cycle. Also provide the rationale for these recommendations.
- Look beyond the next cycle when recommending assessment priorities (a three-cycle horizon?).

Stock Assessment Reviews

- Attempt to schedule only 2 full assessment reviews per STAR panel.
- Schedule earlier reviews of more contentious or complicated assessments.
- Schedule earlier review of “benchmark” data-poor assessments to serve as a guide on how to resolve common problems when reviewing such assessments.
- Continue to distribute STAR panels coastwide.
- Emphasize recommended improvements from previous STAR panel reports in the review.
- More discipline needed in reviews to ensure assessments comply with the Terms of Reference (i.e., updates need to comply by not entertaining new models).
- Continue to have one reviewer attend all STAR panels to provide continuity. Consider a second “continuous” reviewer.
- Provide a LAN and a printer at each STAR panel meeting.
- Rapporteurs should produce real-time assessment summary tables to track changes in the assessment as the review progresses, STAR requests, and STAT responses to those requests. A summary table of important parameters, such as the recruitment time series by species, should be provided. Such summaries should be made available on a web site to disseminate information to other stock assessment teams, STAR panel reviewers, and other advisors to the process.
- Pay more attention to the use of priors in assessments.
- Continue to have SSC members chair STAR panels, but rotate the chair assignments through the stock assessment reviews.
- Formalize the roles of GMT and GAP representatives at STAR panels.
- Updated assessments should only be reviewed by the SSC.

SSC Reviews

- Stagger the SSC reviews (and Council adoption step) across more meetings.
- Improve the management advice in SSC statements recommending assessments. The GMT, GAP, and Council require more specific advice on how assessment results should be applied to management decision-making.
- Schedule Council debriefings with stock assessment lead authors and STAR chairs.
- Consider a joint SSC-GMT-GAP review rather than an SSC review in isolation.
- SSC members who chaired a STAR panel need to recuse themselves when voting to recommend or reject an assessment. Consider assigning an SSC member other than the one who chaired the STAR panel to lead the discussion on an assessment.

Terms of Reference

- Include a requirement for STAR panels to specify whether an assessment is sufficiently developed to do an update in the next cycle.
- Mandate two types of decision tables in assessments- one to address model uncertainty and to portray relative risk of adopting results of alternative models for management decision-making; and one to address statistical uncertainty within the base model.
- Mandate complete executive summaries and decision tables in pre-STAR draft assessments.
- Require an explicit section in each assessment responding to recommendations for improvement from previous STAR panels.
- Require a review of available information from Canadian and Alaskan assessments in each assessment (for stocks with a northerly trans-boundary distribution).
- Require inclusion of maps depicting the scope of the assessment in each assessment.
- Consider adopting a process where rules are established regarding how and who decides if an omission is critical enough to reject an assessment.
- Consider multi-species assessments and/or data reviews in the process.
- Mandate that SSC members who chaired a STAR panel need to recuse themselves when voting to recommend or reject an assessment.
- Adopt a final stock assessment Terms of Reference in April 2006, but defer adoption of the rebuilding analysis Terms of Reference until after MSA re-authorization, resolution of NS1 guidelines, and/or court rulings on rebuilding plans.

PFMC

02/03/06

F:\master\stk\Groundfish\Review\STAR05\Review Workshop\StockAssessmentRevWorkshop_Jan1306_Minutes.doc