STAR Panel Report on the Stock Assessment of Pacific Hake (Whiting) in U.S. and Canadian Waters in 2003.

Overview

On February 2nd-4th a joint Canada-US Pacific Hake STAR Panel met in Seattle, WA to review the stock assessment by Helser et al (2004). The Panel operated according to terms of reference for STAR Panels, but the Panel attempted to adhere to the spirit of the Canada-US Pacific Hake treaty. Both a Panel member and advisor from Canada participated in the review (see List of Attendees). The revised stock assessment and the STAR Panel review will be forwarded to the Pacific Fishery Management Council, council advisory groups, and to Canadian DFO managers and the PSARC Groundfish Sub-committee. The STAT Team was represented at the meeting by Thomas Helser, Richard Methot and Guy Fleischer. Public questions were entertained during the meeting.

The STAR Panel members received a draft of the assessment 10 days prior to the meeting, which was sufficient time to adequately review the assessment. The meeting commenced on February 2, 2004 with introductions followed by a presentation by Guy Fleischer covering the 2003 acoustic survey. After the acoustic survey presentation, Tom Helser presented a detailed description of the stock assessment. Panel discussion continued until the meeting was adjourned February 4th, 2004.

The 2004 assessment used an age structured assessment model developed in AD Model Builder similar to Dorn et al (1998). Major differences between the 2003 assessment and the 2001 assessment included: 2003 acoustic biomass and age composition data, 2002 – 2003 fishery age composition and catch, and 2002-2003 SWFSC juvenile survey data. The initial age structure in the model was extended back from 1973 to 1966. Deep water and northern expansion factors calculated from the 1995-2001 acoustic surveys were applied to the 1977-1992 surveys to correct for restricted survey areas in those years. The new expansion factors were based on age-specific hake distribution in later surveys with adequate spatial coverage and took into account the different geographic distribution of the stock under El Nino, La Nina, and typical years.

While there is room for improvement in the assessment model, as detailed in our recommendations, the Panel generally concurred with methods used in the assessment. The STAR Panel considered the stock assessment document complete and suitable for use by the Council and advisory bodies for ABC projections. The two models carried forward for ABC projections were defined by differences in assumed acoustic survey catchability (q=0.6 and 1.0) and were intended to serve as plausible lower and upper bounds on stock status. STAR Panel voted to view both models (q=1, and 0.6) as equally likely, but the decision was a compromise, with some diversity of opinion among Panel members (see Areas of Disagreement). The STAR Panel commends the STAT team for a well-written document and especially for their cooperation in performing the many additional analyses requested during the meeting (see list of new analyses requested by the STAR Panel).

Summary of stock status

Our understanding of the level of abundance of Pacific hake was changed by this assessment, although the pattern of the stock trajectory is very similar to past assessments. The previous hake assessment in 2002, estimated spawning stock size to be at 20% of unfished in 2001. Because the stock was estimated to be below B25%, Pacific hake were declared overfished in 2001. New information in the 2003 assessment includes fishery age composition in 2002 and 2003, but more importantly, the results of the 2003 acoustic survey. The increase in biomass in the 2003 acoustic survey and the dominance of the 1999 year class in both fishery and survey data suggest that the 1999 year class is even higher than previously estimated. The revised northern expansion factors derived from surveys in 1995-2001 suggested that biomass was higher in earlier years of the modeled time period. In addition, changes in the model structure accounted for some of the estimated increase in biomass in those earlier years. These changes produced a fairly significant difference between the current assessment and previous assessments, though estimates of overall trend and current stock depletion were robust (Figure 1).

Stock size in 2003 was estimated to be 2.7 to 4.2 million t. for models with fixed acoustic survey q= 1.0 and 0.6, respectively. A q =1.0 implied that the acoustic survey produces an estimate of absolute biomass, while a q= 0.6 implied that the acoustic survey biomass estimate were on average lower than stock biomass. Both model scenarios allowed dome-shaped selectivity for the acoustic survey, thus allowing for even lower effective q levels for young and old fish.

Stock depletion in 2003 was estimated to be 47% of unfished for a model with an acoustic survey q =1.0 and 51% of unfished for a q= 0.6. Estimates of stock depletion in 2001 in the current assessment ranged from 27-31% of unfished, indicating that the stock approached, but did not drop below the B25% overfished threshold. Under both assumptions of catchability, the stock has rebuilt to levels above B40% in 2003.

Mature biomass was projected to decline from 2004 to 2007 to below 30% of unfished due to the absence of strong recruitment after the 1999 year class. A sharp increase followed by a gradual declined is a typical pattern of biomass variability for a stock with highly variable recruitment. Lower harvest rates would lessen the projected stock decline but not reverse it.

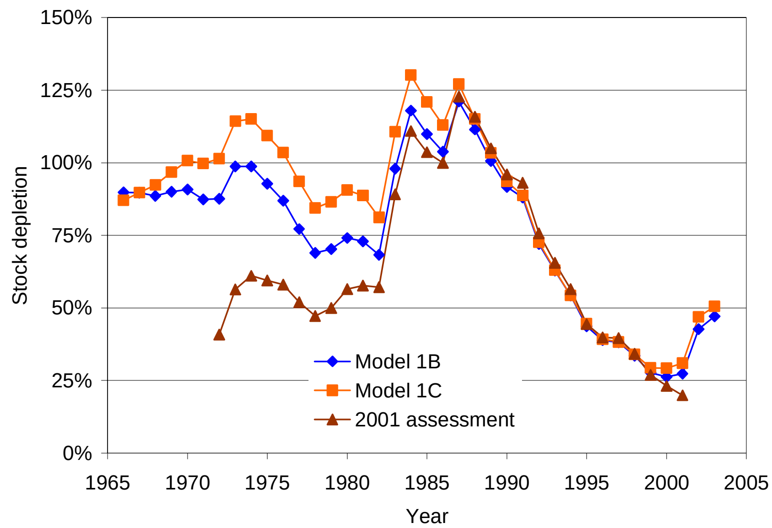

Figure 1. Estimates of stock depletion in the 2003 and 2001 Pacific hake stock assessments.

List of New Analyses Requested by the STAR Panel

A number of new analyses and model runs were requested by the STAR Panel and completed at the meeting by the STAT team. The following list describes each request, followed by the reason for the request and outcomes of the analysis.

Request: The Panel requested that the STAT team use the Jolly-Hampton method to calculate sampling CVs for the 2003 acoustic survey biomass. Reason: to get a better estimate of the sampling variability of the acoustic survey. Outcome: Post-stratification estimates of sampling CV = 0.37.

Request: The STAR Panel requested that model option 5 (from the draft assessment document) be structured to estimate acoustic survey selectivity in 3 periods. Each period grouped as years consisting of El Nino, La Nina and all other years. This model was referred to option 5a. Reason: More objective method to deal with changes in acoustic selectivity. Outcome: Small improvement in fit to acoustic age-composition from separating out only the El Nino years, but not all 3 (El Nino, La Nina and others).

Request: The STAR Panel requested that option 5 (from the draft assessment document) be structured to increase the acoustic survey CVs to 0.5 from 1977-1989 and O.1 from 1992-2003 (Model 5b). Reason: more realistic estimates of uncertainty for those early years.

Request: The STAR Panel requested that Option 5 (from the draft assessment document) be structured to increase the acoustic survey CVs to 0.5 from 1977-1989 and 0.3 from 1992-2003 (Model 5c). Reason: The assumed CVs were too small given the errors associated with survey method. Outcome: The inflated CVs were more internally consistent, as measured by mean square error in model fit to the acoustic survey.

Request: The STAR Panel requested that option 5 (from the draft assessment document) be structured to remove 1986 acoustic biomass and age composition data (Model 5d). Reason: The change in signal strength in pre- and post survey calibrations makes the 1986 estimate highly questionable and potentially biased. Outcome: Small changes to the model results.

Request: The STAR Panel requested that Option 4 (from the draft assessment document) be restructured to include a single acoustic selectivity, a biomass CV=0.5 for years 1977-1989 and CV=0.3 for years 1992-2003, and acoustic *q* estimated (model 4a). Reason: Explore the possibility of using an estimated acoustic *q* instead of fixed. Outcome: the estimates of *q* are low, but the fit to the survey are better.

Request: The STAR Panel requested that Option 4 (from the draft assessment document) be structured to remove 1986 acoustic biomass and age compositions (Model 4b). Reason: Calibration problems in the acoustic gear make the 1986 estimate potentially biased. Outcome: the estimate for q was slightly higher than for Model 4a, but was still unrealistically low.

Request: The STAR Panel requested that Option 4 (from the draft assessment document) be structured to estimate acoustic *q*, assume full selectivity for all ages, and remove survey age composition data and, the 1986 acoustic survey biomass. Reason: The age-composition of the acoustic survey may not be representative of the acoustic gear. Outcome: the estimate for q dropped relative to Model 4b and the overall fit to the data deteriorated.

Request: The STAR Panel requested that Option 4 (from the draft assessment document) be structured to assign effective sample sizes to the US fishery age composition = 400 (Model 6a). Reason: Increasing the effective weight on the age compositions was done to determine if that resulted in a directional movement of the standard deviation of the standardized residuals. Outcome: Increasing the effective sample sizes resulted in an increase in the standardized residuals and the STAR Panel agreed to adjust the effective weights on the other age composition datasets.

Request: The STAR Panel requested that Option 4 (from the draft assessment document) be structured to replace the random walk in the fishery age selectivity with 3 periods of constant selectivity (1966-79, 80-89, and 90-2003) based on changes in the fishery. Reason: To determine if a reduction in the number of parameters resulted in degradation in fits to the age-composition data. Outcome: Problems arose in the age-composition fits when the random walk was removed that were deemed unacceptable.

Request: The STAR Panel requested a model structured (model 1a) as acoustic survey *q*=1, adjust input variances in the model to be consistent for all age compositions (300 US and 130 Canada commercial fisheries and 60 for acoustic), survey acoustic biomass CV= 0.5 (1977-1989) and CV=0.3 (1992-2003), random walk in commercial fisheries selectivity and remove the 1986 acoustic data (biomass and age composition). The SWFSC midwater juvenile survey CV=1.1. Reason: The STAR Panel wished to produce an internally consistent model based on the weightings that preserved the historical use of acoustic *q* =1, and assigned realistic CVs to both the acoustic biomass estimates and the SWFSC juvenile index. Outcome: this model did not provide good fits to the 1977-1989 survey biomass estimates.

Request: The STAR Panel requested a model structured (model 1b) as model 1a but with CV= 0.2 (1977-1989) and CV=0.1 (1992-2003). Reason: The STAR Panel wished to produce an internally consistent model based on the weightings that preserved the historical use of acoustic *q* =1, with lower CVs on the acoustic biomass series so that the model would follow the same trend as the acoustic biomass estimates. Outcome: this model provided the best fits to the survey acoustic biomass indices with the q=1 assumption.

Request: The STAR Panel requested a model structured (model 2a) to estimate acoustic *q*, acoustic biomass CV=0.5 (1977-1989) and CV=0.3 (1992-2003), remove 1986 acoustic data (biomass and age composition), the acoustic survey age composition is decoupled from the survey biomass and a uniform selectivity is imposed on the acoustic survey. The SWFSC midwater juvenile survey CV=1.1. Reason: The STAR Panel wished to produce an internally consistent model based on the weightings that estimated *q*, and gave realistic CV to the acoustic biomass. Outcome: The estimate of acoustic survey was *q* =0.28, which was considered implausible by the Panel

Request: The STAR Panel requested a change to the above request that structured the model (Model 2b) to estimate acoustic *q*, acoustic biomass CV=0.5 (1977-1989) and CV=0.3 (1992-2003), remove 1986 acoustic data (biomass and age composition), the acoustic survey age composition was removed and a uniform selectivity is imposed on the acoustic survey. The SWFSC midwater juvenile survey CV=1.1. Reason: The STAR Panel wished to produce an internally consistent model based on the weightings that estimated *q*, and gave realistic CV to the acoustic biomass but without the acoustic age composition data. Outcome: The estimate of acoustic survey was *q* =0.21, and that was agreed upon by all participants as unrealistically low.

Request: The STAR Panel requested a model with acoustic survey *q* =0.6(model 1c), tune the input variances in the model to be consistent for all age compositions (300 US and 130 Canada commercial fisheries and 60 for acoustic), and survey acoustic biomass CV= 0.5 (1977-1989) and CV=0.3 (1992-2003), random walk in commercial fisheries selectivity and remove the 1986 acoustic data (biomass and age composition). The SWFSC midwater juvenile survey CV=1.1. Reason: Establish a upper bound of stock status. Outcome: This run provided improved fit to the acoustic biomass survey indices.

Request: Panel and STAT team agree to use Model 1b and 1c to provide a range bounding the knowledge of stock status.

Request: The Panel requests the STAT team do projections using MCMC output from both Model 1b and 1c using both F40% and F45% harvest rates.

Request: The Panel requested that the STAT team provide a decision table using the different Models as states of nature and the F40% and F45% harvest rates as management decisions.

Request: The STAR Panel requested that standard deviation of the standardized residuals be calculated for each data source. Reason: A diagnostic of model fit.

Technical merits and deficiencies

Acoustic survey

The acoustic-trawl survey data were used in the assessment to provide biomass indices and estimates of proportion at age. The surveys are triennial from 1977 to 2001, with the latest survey in 2003. The surveys from 1977 to 1989 cover a smaller depth range than the later surveys and the 1977 to 1992 surveys do not go as far north as the later surveys. Deep water and northern expansion factors were applied to the appropriate surveys in an attempt to make the whole time series consistent.

The survey design appeared to have been relatively consistent from year to year (with the exceptions of coverage). Transects were typically east to west generally running between 50 m and 1500 m depth contours. Transects were allowed to be extended to deeper water if fish densities were high near the normal stopping point. Transects were done during the day with most trawling during the day for target identification and collection of biological samples.

Merits of the time series include:

- The survey area covers a very large proportion of the adult hake distribution (i.e., areal availability is near to 1).
- Hake form large (mainly) midwater aggregations during the time of the survey so marks are easily identified and there is limited undersampling in the “dead zone” near the bottom.
- Sampling intensity was generally good with 80-100 transects in all years.

There were also some important considerations when the survey data are used in stock assessment models:

- The length target strength relationship for hake is based on a small number of in situ measurements. These were made during the night (from low density marks) when tilt angle distributions and swimmbladder inflation levels could differ from those during the day in high density marks (where most the biomass is found). Thus there is the potential for significant bias in the indices when they are used as *absolute* abundance.

- No routine calculations of variance are made for the survey estimates (biomass or proportions at age). The variance assumptions made in an assessment model must therefore be based on model residual patterns and somewhat arbitrary decisions.
- The proportion at age data are derived from the target identification trawls. These are necessarily *targeted* on marks seen on the acoustic transects. It is not clear that the resulting age samples are representative of the population. However, the triennial bottom trawl survey age frequencies are very similar to the acoustic survey age frequencies which does suggest the survey samples are representative.
- The precision of the biomass indices will vary from survey to survey.
- The precision of the proportions at age will vary from survey to survey and will have a complex error structure.
- The biomass indices are correlated with the estimates of proportion at age (in a complicated way).
- The pre and post survey calibration constants for the 1986 survey differed by a significant factor. Application of the post-cruise value would have led to a 48% increase in the hake biomass. In previous assessments, the conservative estimate was used. The reasons for this are not entirely clear to the current Panel and all 1986 survey data were removed from the final runs.
- The 2003 acoustic survey used the W.E. Ricker for the entire survey. Earlier surveys used the Miller Freeman for the U.S. portion of the survey (1995-2001), or for the entire survey (1977-1992). Inter-vessel comparisons between the Miller Freeman and the W.E. Ricker during previous acoustic surveys have not found large differences in the summed acoustic backscatter along transects between the two vessels, though the power to detect moderate differences is low.
- Midway through the 2003 survey, it was found that the face of the transducer on the W.E. Ricker was encrusted with barnacles. Based on calibrations before and after their removal, the signal loss due to biofouling was 0.61 dB, implying a change of acoustic backscatter of ~30%. This signal loss was corrected for, but additional uncertainty is associated with that portion of the 2003 biomass estimate as a result of this correction.

Catch and catch at age

Total catch was available from 1966-2003 by nation and fishery. The accuracy of the total catch estimates was not considered by the Panel, but they are believed to be accurate from 1977-2003. In the earlier period the total catch may have been underestimated.

There has been extensive sampling of the commercial catch, with catch at age estimates for the U.S. fishery from 1973-2003 and for the Canadian fishery from 1977-2003. Some adjustments for ageing error were made to these data by accumulating numbers at age for some cohorts in some years. The Panel did not consider these specific adjustments or the question of ageing error in general. However, plots of the estimated proportions at age very clearly show the progression of strong cohorts (so ageing error is perhaps a minor issue).

Estimates of variance for the proportions at age data are not reported in the assessment. As with the acoustic data this requires that model assumptions with regard to variance be based on model output and somewhat arbitrary decisions.

Recruitment indices

The SWFSC midwater trawl survey targeting pelagic juvenile rockfish was used to provide a recruitment index from 1983-2003. This survey covers a small geographic area relative to the distribution of the juvenile hake. However, the indices have been shown to have a significant correlation to model estimates of recruitment.

Differences between this time series and a shorter recruitment time series over a wider area (PWCC-NMFS midwater trawl survey) were noted. It is not clear whether the two time series are contradictory (as they are both very imprecise). The Panel did not consider whether it was appropriate to include the PWCC-NMFS indices in the assessment runs.

Biological parameters

Year specific weights at age were used in all years for each fishery and survey because of significant variation in the observed weight at age. A constant and age independent estimate of natural mortality was used. A constant female maturity at age vector was also used. The Panel did not consider the derivation or use of these estimates in any detail.

Stock assessment model and estimation procedure

The single-sex age structured model uses standard population dynamics equations. The Canadian and U.S. fisheries are modeled as distinct year-round fisheries. Fishing selectivity patterns are year specific (constrained by a random walk) to allow for changes in fleet composition and shifts of fish distribution (across the border). The acoustic time series is modeled using a single selectivity pattern which applies to both the biomass indices and the estimated proportions at age.

The estimation procedure is essentially maximum likelihood with Bayesian extensions for estimating parameter uncertainty. The initial runs presented to the Panel all assumed the acoustic biomass indices were absolute (acoustic catchability, $q = 1$). This assumption has been made for all previous hake assessments (although it was questioned in the 2002 STAR Panel meeting). Runs where q was estimated in the model had been done (but not presented) and they suggested values of q substantially less than 1 (and consequently much higher biomass).

The Panel supported the use of the general modeling and estimation procedure but had concerns about some aspects of the approach. The major concern was the assumption of *q* = 1. It was suggested that the alternative approach of estimating *q* should be more fully explored with results being presented to the meeting. The STAT Team presented the results of approximately 20 runs which included several with *q* freely estimated. After exploration of residual patterns and diagnostic statistics from these runs, the meeting selected four runs for further evaluation: models 1a, 1b (*q* = 1); and runs 2a, 2b (q estimated).

Run 1a used variance weightings assumptions that were consistent with model residuals (standard deviation of standardized residuals near to 1). However, the predicted biomass from the run was substantially higher than the observed biomass for the early part of the acoustic time series (1977-1989). In run 1b the acoustic biomass indices were given greater weight to encourage a better fit to the early part of the time series (consistent with the assumption of *q* = 1 over all years). This did improve the fit to the time series, but made the model residuals inconsistent with the variance assumption. Nevertheless it was considered by the meeting to represent the best model run based on the assumption of *q* = 1.

Runs 2a and 2b gave reasonable fits to the whole acoustic time series but produced estimates of *q* which were considered implausibly low. This judgment was made after deriving plausible lower and upper bounds for *q* (0.55-1.25) based on four factors: a real availability, vertical availability, target identification, and target strength. In order to provide a credible alternative to run 1b, the meeting adopted run 1c which had a fixed *q* = 0.6 (a variation of run 1a). This provides a credible value of *q* without unduly compromising the fit to the data.

Neither of runs 1b and 1c was entirely satisfactory. Each derives from an approach which has been compromised to some extent in order to achieve credible results (a better fit in one case and an acceptable value of *q* in the other). The meeting considered that the best approach to use in the future is to develop a Bayesian prior on *q*. There was insufficient time to do that during this meeting.

Harvest policy

The Panel did not address the issue of appropriate harvest rates for Pacific hake. The recent US-Canada hake agreement specifies F40% with an 40-10 adjustment as the default harvest rate. The harvest policy review Panel also recommended an F40% harvest rate, citing a meta-analysis by Dorn et al (1998) of hake stocks world-wide that suggested harvest rates in the F40%-F45% range could be considered appropriate proxies for FMSY, depending on the level of risk aversion. The high recruitment variability of hake results in rapid increases and subsequent declines in abundance and yield. Based on stock projections, it is apparent that the stock may decline to near or below the depleted threshold (25% unfished) without the recommended harvest rate ever being exceeded. We concur with the assessment authors that a new examination of the harvest policy that takes into account this variability is needed for this highly fluctuating stock.

Areas of Major Uncertainty

While there is uncertainty in both data and the model structure, the Panel concluded that the major source of uncertainty lies in the assumption of acoustic survey *q*. The STAR Panel and STAT Team attempted to estimate acoustic *q*, but the resulting estimates were deemed unlikely by all participants and were therefore not brought forward. An ad hoc approach of determining bounds to *q* was developed based on expert opinion about the magnitude of error in the major sources of uncertainty in acoustic surveys (Table 1). Those bounds ranged *q*=0.55- 1.3. The Panel and STAT team concluded that q>1 was unlikely and thus bounded uncertainty using *q*=0.6 and q=1. The Panel and STAT team concluded that we did not have sufficient information at the meeting to determine *q* more precisely.

Table 1. Upper and lower bound on the uncertainty in estimated biomass from selected components of the acoustic survey. Upper and lower bounds of acoustic *q* are a product of these values

Source of Error	Lower bound	Upper bound
Area availability	0.95	1
Vertical availability	0.8	0.95
Target identification	0.9	1.1
Target strength	0.8	1.2

Areas of Disagreement

The only source of disagreement among STAR Panel members and STAT team members was the appropriate weighting to give the model that fixed *q*=1 and *q*=0.6. The Panel and STAT team agreed that *q* was uncertain and viewed the equal weightings on both models as a compromise. We note here, however, that some Panel members supported higher weighting on the *q*=1 scenario, and other members preferred a higher weighting on *q*=0.6.

Research Recommendations:

General recommendation for data: all data (primarily the acoustic survey data, fishery and age composition data) used in the assessment should be critically evaluated. Data that are determined to be biased or suspect should be dropped. For example, the consistency of the ageing data between years and between laboratories should be reviewed. This work should be fully documented so that the reasoning for the decisions is preserved.

1. Acoustic survey recommendations:

a. Determine whether there are differences in survey performance between the WE Ricker & Miller Freeman. These include differences in mid-water and bottom trawl efficiency as well as differences in acoustic capabilities between the vessels. Analyze the available data to determine if we can continue to accept the null hypothesis that there is no difference in survey performance between these vessels.

b. Perform a detailed meta-analysis across all survey years: compare spatial distributions of hake across all years and between bottom trawl and acoustic surveys to estimate changes in catchability/availability across years.

c. Generate appropriate estimates of variability for every survey year.

d. Review the methods used to estimate proportions at age for the acoustic survey with particular regard to the representativeness of trawl samples.

2. Estimation of target strength:

a. Evaluate the current target strength for possible biases, particularly the use of nighttime experiments which are applied to daytime survey transects. Explore alternative methods for estimating target strength.

b. Assess the value of the recent Canadian hake target strength observations and, if these are assessed to be useable, add these into the target strength model.

c. Commission the acquisition of additional in-situ observations to increase the model sample size.

3. Model enhancements:

a. Add in bias correction for log-normal distribution in appropriate likelihoods.

b. Recode the model so that projections are done as a post-MCMC procedure.

c. Develop an informed prior for the acoustic *q*. This prior should be used in the model when estimating the q parameter

d. Consider the development of a sex-structured model.

e. Investigate alternative methods to model annual variability in fishery selectivity. Identify the covariates that influence fishery selectivity.

f. Investigate the interaction of the dome-shaped selectivity functions with the fixed value of M. This investigation should include determining whether there is a trade-off between M and the declining limb of the selectivity function. Investigate the possibility of age-specific M.

g. Investigate alternatives to applying a single estimated acoustic selectivity based on trawl samples to the acoustic biomass indices.

4. The STAR Panel had difficulty completing its assigned task during a three day review. At least a full week is needed for a more thorough review of the input data and the assessment model.

List of Attendees

STAT team members present
Thomas Helser (Northwest Fisheries Science Center)
Richard Methot (NMFS Office of Science and Technology)
Guy Fleischer (Northwest Fisheries Science Center)

Star Panel members:
Kevin Piner, (chair, NOAA/NMFS, Southwest Fisheries Science Center)
Jeff Fargo, (Fisheries and Oceans, Science Branch)
Martin Dorn, (SSC, NOAA, NMFS, Alaska Fisheries Science Center)
Patrick Cordue, (CIE reviewer)

STAR Panel Advisors:
Xi He, (GMT, NOAAA/NMFS, SWFSC)
Rod Moore, (GAP, West Coast Seafood Processors)
Paul Starr (Industry Consultant Canada, Canadian Groundfish Research and Conservation Society)

List of Public Attendees
Vera Agostini (University of Wahington)
Ian Stewart (University of Washington)
Vidar Wespestad (Pacific Whitting Conservation Cooperative)
Rick Dunn (Trident)
Owen Hamel (NOAA/NMFS, Northwest Fisheries Science Center)
Steve Joner (Makah Tribe)
Mike Buston (Leader Fishing)